

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2004, 23:30:45 ; Search time 529 Seconds

(without alignments)
2589,979 Million cell updates/sec

Title: US-10-043-649-2

Perfect score: 1351

Sequence: 1 MGSIPSRKSLPSPSLSSV.....RESISFYISUNDEAVSDDA 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 413486 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O=/cgm2.1/USPTO.spool/US10043649/runat.16112004.060535.28709/app.query.fasta_1.455
-DB=N Geneseg 23Sep04 -QWMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOEXT=0 -UNITS=dits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=Pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10043649 @CGN 1.1 470 @runat.16112004.060535.28709 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEX=7

Database : N Geneseg 23Sep04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1351	100.0	786	6	AAL44089 Human mod
2	1351	100.0	786	6	ABQ74343 Human Src
3	1351	100.0	2567	6	AAD43980 Human Src
4	1351	100.0	3646	10	ADF90741 Human hep
5	1347	99.7	1183	6	ABK61465 Human CDN
6	1345	99.6	2788	10	ADF82458 Leukaemia

7	1273	94.2	837	3	AAC77202 Human ORF
8	1210.5	89.6	1413	6	ABQ9374 Human cod
9	1200.5	88.9	737	6	AAL44090 Mouse MAR
10	1132.5	83.8	2049	5	AAS74750 DNA encod
11	1032	76.4	1348	6	AAL44087 Mouse mod
12	830	61.4	763	6	ABQ98670 Human ORF
13	731	54.1	603	5	AAS74748 DNA encod
14	647	47.9	864	5	ADL63090 Human ova
15	586	43.4	875	6	ABQ99151 Human ORF
16	492	36.4	3756	8	ABX62975 Human act
17	491	36.3	2109	4	AAS02049 DNA encod
18	491	36.3	2665	6	ABL65189 Human cand
19	491	36.3	2665	6	ABK83738 Human CDN
20	491	36.3	2665	8	ACC81091 Human Src
21	491	36.3	2665	12	ADL83115 Human PRO
22	491	36.3	2665	12	ADP12833 Reference
23	491	36.3	3452	12	ADQ22520 Human sof
24	452.5	33.5	444	6	ABQ98669 Human ORF
25	374.5	27.7	2298	6	ABK83935 Human CDN
26	374.5	27.7	2298	10	ADD19015 Human dis
27	374.5	27.7	2298	10	ADP81586 Leukaemia
28	374.5	27.7	2298	11	ADL1832 Human CDN
29	374.5	27.7	2298	11	ADN95431 Human BEC
30	374.5	27.7	2298	12	ADL22891 Human MP2
31	374.5	27.7	2298	12	ADN04497 Antisort
32	374.5	27.7	2298	10	ADP10425 Reference
33	374.5	27.7	4175	10	ADP81585 Leukaemia
34	364.5	27.0	1924	12	ADL04091 Human CCK
35	364.5	27.0	1926	6	ABK83940 Human CDN
36	364.5	27.0	1926	12	ADJ71658 Human NOV
37	364.5	27.0	2015	6	ABL66673 Lung cand
38	364.5	27.0	2015	6	ABK83939 Human CDN
39	364.5	27.0	2015	10	AAD62155 Human hae
40	364.5	27.0	2015	11	ADL131779 Human CDN
41	364.5	27.0	2015	12	ADL22888 Human MP2
42	364.5	27.0	2015	12	ADP12819 Reference
43	364.5	27.0	2025	11	ADL17038 Gene enco
44	364.5	27.0	2341	10	ADB53793 Human pro
45	362.5	26.8	1911	6	ABK63704 Rat seque

ALIGNMENTS

RESULT 1
ID AAL44089 standard; CDNA; 786 BP.
XX
AC AAL44089;
DT 03-OCT-2002 (first entry)
DE Human modulator of antigen receptor signalling protein coding sequence.
KW Human; gene; ss; gene therapy; modulator of antigen receptor signalling;
KW MARS; tumour suppressor gene; Src-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..786
FT /tag= a
FT /product= "Human MARS protein"
XX
PN W0200242452-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-CA001662.
XX
PR 27-NOV-2000; 2000CA-02324663.
XX

PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
 XX
 XX Mcglade JC, Loreto MP;
 PI
 XX
 DR WPI; 2002-566564/60.
 DR P-PSDB; AAO15457.
 XX
 PT New isolated modulator of antigen receptor signaling protein or its
 PT fragment, useful for treating malignant disorders such as myeloid
 PT malignancies, autoimmune disorders and myeloproliferative disorders.
 PS
 PS Claim 12; Page 75, 110pp; English.
 CC The invention comprises the amino acid and coding sequences of modulator
 CC of antigen receptor signaling (MARS) proteins. The MARS protein is a
 CC putative tumour suppressor gene and exhibits structural and sequence
 CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and
 CC protein sequences of the invention are useful for the treatment of
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 CC disorders, immunosuppression, myeloproliferative disorders and
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.
 CC breast cancer). The present cDNA sequence encodes a human MARS protein
 XX
 XX Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 1,36e-110 Length: 786
 Score: 1351.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-10-043-649-2 (1-261) x AAL44089 (1-786)
 QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerSerVal 20
 DB 1 ATGGGAAGTCTGCCACAGAGAAATCTCTGCAAGCCCAAGCTTGAGTCTCTCTGTC 60
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 DB 61 CAAGGCCAGGAGACTGTGACCATGAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 DB 121 GGCAGTTTCCCGGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
 QY 61 IleValSerGluAspGlyAspTTPTRPThValLeuSerGluValSerGlyArgGluTyr 80
 DB 181 ATCGTCTGAGGATGAGACTGTGTGACGCTGTCTGAGTCTGACGACAGAGATAT 240
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrPheUtyrGluLysSer 100
 DB 241 AACATCCCCAGCGCTCCACGTGGCCMAAGTCTCCCATGGGCTGTAATGAGGGCTGAGC 300
 QY 101 ArgGluLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 DB 301 AGGAG 360
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 DB 361 CGGAGAGCCAGACCGAG 420
 QY 141 SerTyrAspArgIleArgHisTyrArgGlyLeuHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 DB 421 TCTCTGGAGCCGGATTCACACACTACAGGATCCACTGCTTACCAATGGCTGGCTGATC 480
 QY 161 SerProArgGluThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 DB 481 TCACCGGAGCTCACTCCCTCAGCTCCAGAGCCGTGGTGAACCATTACTTGAGCTGGGG 540
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200

DB 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTCTGACAGAGGCTGGCCCTCCCT 600
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLeu 220
 DB 601 GGCAGAGATATACCCCTACTGTGACTGTGACAGAGACACACACTCAACTGAGAGAGAGCTG 660
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluLeuSerLeuSerGluGly 240
 DB 661 GACAGCTCCCTCTGTTTCTGTAAGCTCCACAGGGAGAGAGTCTTCTTCAGTGAAGGT 720
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
 DB 721 CTCGGGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGAGCTGTCTTGTGATGAT 780
 QY 261 Ala 261
 DB 781 GCC 783
 RESULT 2
 ABQ74343
 ID ABQ74343 standard; cDNA; 786 BP.
 XX
 XX ABQ74343;
 XX
 DT 15-OCT-2002 (first entry)
 XX
 DE Human Src-like inhibitory molecule (SLIM) encoding cDNA.
 XX
 KW Human; Src-like inhibitory molecule; SLIM; Src-like adaptor protein;
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;
 KW gene; ss.
 XX
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 1..786
 FT CDS /*tag= a
 FT /product= "SLIM"
 FT /note= "Src-like inhibitory molecule"
 XX
 XX MO200255707-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 10-JAN-2002; 2002WO-US000718.
 XX
 PR 10-JAN-2001; 2001US-0260953P.
 XX
 PA (RIGE-) RIGEL PHARM INC.
 XX
 PI Holland SJ, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y,
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N,
 XX
 XX WPI; 2002-575432/61.
 DR P-PSDB; ABP52187.
 XX
 PT New src-like inhibitory molecule protein, useful for treating
 PT immunodeficiency disorders and inflammatory disorders, comprises N-
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.
 PS
 PS Claim 3; Fig 2a; 91pp; English.
 CC The present sequence encodes the human Src-like inhibitory molecule
 CC (SLIM) protein (I). The present invention describes a SLIM protein
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and
 CC anti-HIV activities, and can be used as a modulator of lymphocyte
 CC activation, and of ubiquitination of a Cbl target protein, and in gene

CC inflammation disorders, diseases and conditions, rheumatoid arthritis, CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's CC and ulcerative colitis), allergies, particularly those involving CC hyperactivity of B-cells and T-cells, or other immune cells, such as CC mast cells or eosinophils, autoimmune diseases such as systemic lupus erythematosus and multiple sclerosis, pulmonary diseases including CC asthma, acute respiratory distress syndrome, and chronic obstructive CC pulmonary disorder, tissue/organ rejection and cancer. The invention is CC useful in gene therapy. The present sequence is human SLAP-2 CDNA

XX Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6,29e-110	Length:	2567
Score:	1351.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-043-649-2 (1-261) x AAD43980 (1-2567)

```

QY      1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
DB      415 ATGGGAAGTCTGCCACAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 474

QY      21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB      475 CAAGGCCAGGAGCCTGTGACCATGAGACAGAGAAAGCAAGGCCACAGCCGTG6CCCTG 534

QY      41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB      535 GGCAGATTCCCGGAGGTGCGCCGCGAGCTGTGCTGAGACTCGGAGGAGCATTGACC 594

QY      61 IleValSerGluSpGlyAspTPTPTThrValLeuSerGluValSerGlyValArgGluTyr 80
DB      595 ATCGTCTCTAGATGAGATGAGACTGCTGAGCCGCTGTCTGAAGCTCAGGACAGAGATAT 654

QY      81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGlnGlyLeuSer 100
DB      655 AACATCCCAAGCGTCCACGCTGGCCAAAGTCTCCATGGTGGCTGTATGAGGGCTTAGC 714

QY      101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB      715 AGGAGAAACAGAGAGAACTGCTGTGTACTGTGGAACCTGGAGGGGCTTCTCCTATC 774

QY      121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB      775 CGGAGAGCCAGACACAGAGAGGCTTCTTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 834

QY      141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160
DB      835 TCCGGSAGCCGAGTACACACTACAGATCACCGCTTGACAAATGGCTGGCTGTATATC 894

QY      161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
DB      895 TCACCGGCTCAGCTCCCTCCCTCCTCAGGCTGAGGCTGAGCCATTACTGAGACTGGCG 954

QY      181 AspAspIleCysCysLeuLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB      955 GATGACATCTGCTGCTCTCACTCAAGAGCCCTGTGTCTCTGCAAGAGGCTGGCTCCT 1014

QY      201 GlnGlyAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrPlyGluLeu 220
DB      1015 GGCAGAGATATACCCCTTACTGTAAGTGTGAGAGACACCACTCACTGGAAGAAGACTG 1074

QY      221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluSerLeuLeuSerGlnGly 240
DB      1075 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGAGGAGGAGTCTTCTTCAAGTGAAGGT 1134

QY      241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnSpGluAlaValSerLeuAspAsp 260
DB      1135 CTCGGGAGTCTCCCTCAGCTTCTTACATCAGCTGAATGACGAGGCTGTCTTTGGATGAT 1194
  
```

QY 261 Ala 261
DB 1195 GCC 1197

RESULT 4

ADPF90741/c
ID ADPF90741 standard; DNA; 3646 BP.

XX ADPF90741;

XX 26-FEB-2004 (first entry)

DE Human hepatic-fibrosis disease marker SEQ ID 203.

KW Hepatic fibrosis; marker; chronic hepatitis; liver cirrhosis;

XX hepatic carcinoma; human; ds.

OS Homo sapiens.

PN JP2003259877-A.

PD 16-SEP-2003.

PF 11-MAR-2002; 2002JP-00065013.

PR 11-MAR-2002; 2002JP-00065013.

PA (SUMITOMO SEIYAKU KK.

XX WPI; 2003-821596/77.

PT Hepatic fibrosis disease markers comprising polynucleotides or

PT antibodies, useful for improved diagnosis, screening and developing drugs

XX to treat hepatitis, to control cirrhosis and carcinoma.

XX Claim 1; SEQ ID NO 203; 313bp; Japanese.

CC The present invention relates to hepatic-fibrosis disease markers
CC (ADPF90539-ADPF90871) and related proteins (ADPF90872-ADPF90917). The
CC sequences are useful for detecting and treating hepatic fibrosis caused
CC by alcohol consumption, virus infection, etc., and the associated chronic
CC hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC precision), so more suitable treatments can be developed and given.

SQ Sequence 3646 BP; 782 A; 954 C; 1031 G; 877 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	9.91e-110	Length:	3646
Score:	1351.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-043-649-2 (1-261) x ADPF90741 (1-3646)

```

QY      1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
DB      3202 ATGGGAAGTCTGCCACAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 3143

QY      21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB      3142 CAAGGCCAGGAGCCTGTGACCATGAGAGCAGAGAAAGCAAGGCCACAGCGTGGCCCTG 3083

QY      41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB      3082 GGCAGATTCCCGGAGGTGCGCCGCGAGCTGTGCTGAGACTCGGAGGAGCATTGACC 3023

QY      61 IleValSerGluSpGlyAspTPTPTThrValLeuSerGluValSerGlyValArgGluTyr 80
DB      3022 ATCGTCTCTAGATGAGATGAGACTGCTGAGCCGCTGTCTGAAGTCTTCAAGGACAGATAT 2963
  
```

QY	81	AsnIleProSerValHisValAlaIaIysValSerHisgIYTrpIleuTYrGIuGIYLeuSer	100
Db	2962	AACATCCTCCAGGCTCACGTGCAGGTCMAAGTCTCCATGGGTGGCTGTATACAGAGGCTTGAGC	2903
QY	101	ArgGIuIysAlaGIuGIuIleuIeuIeuProGIYAsnProGIYGIValaIheIeuIle	120
Db	2902	AGGGAGAAAGCAGAGAGAACTGCTGTGTGTACTGGGAAACCTCGAGGGGCTCTTCATATC	2843
QY	121	ArgGIuSerGIuThrArgArgGIYSerTYrSerLeuSerValArgLeuSerArgProIa	140
Db	2842	CGGGAGACCCAGACCAGAGAGGCTCTACTCTGTGTAGTCCGCTCAGCCGCCCTGCA	2783
QY	141	SerTrpAspArgGIleuArgHisTYrArgIleHisCYLeuAspAsnGIYTrpIleuTYrIle	160
Db	2782	TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGCTGGCTGTATATC	2722
QY	161	SerProArgLeuThrPheProSerIleuGIuIaIeuValaPheHisTYrSerGIuIeuAla	180
Db	2722	TCACCGGCTTCACCTTCCCTCCTCATCTCAGGCGCTGGTGGACCATTACTCTGAGCTGGCG	2663
QY	181	AspAspIleCYCYsLeuIeuIuIesIuProCYsValIeuGIuArgAlaGIYProIeuPro	200
Db	2662	GATGACATCTCTGTGCTACTCAGAGAGCCGTGTGTCGTGAGAGGGCTGGCCGCTCCT	2603
QY	201	GIYIysAspIleProIeuProValIThrValGIuArgThrProIeuAsnTrpIysGIuIeu	220
Db	2602	GGCAGAGATATAACCTTAACCTGTGACTGTGCAGAGGACACCACTCACTGGAAGAAGCTGTG	2543
QY	221	AspSerSerLeuIeuPheSerGIuAlaIaIaThrGIYGIuIuSerLeuIeuSerGIuGIY	240
Db	2542	GACAGCTCCCTCCTGTTTCTGAAAGCTGCCAAGGGGAGGAGCTCTCTTCTCAGTGAAGGT	2483
QY	241	IeuArgIuSerIeuSerPheTYrIleSerIeuAsnAspGIuAlaIaIaSerIeuAspAsp	260
Db	2482	CTCGGGAGGTCCCTCAGCTTCTATCATCAGCTGGAATAGCAGAGGCTGTCTCTTGGATGAT	2423
QY	261	Ala 261	
Db	2422	GCC 2420	

RESULT 5
ABRK61465
ID ABRK61465 standard; cDNA, 1193 BP.
AC ABRK61465;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cDNA encoding protein NOV13.
XX
KW Human; gene: ss; NOVA; gene therapy; cardiomyopathy; atherosclerosis; cell signal processing disorder; metabolic pathway modulation disorder; diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; uterus cancer; immune response; graft-versus-host disease; acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease; hypertension; congenital heart defects; multiple sclerosis; inflammation; Albright hereditary osteodystrophy.
KX KW KM KN
OS Homo sapiens.
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P. 08-FEB-2001; 2001US-0267300P. 20-FEB-2001; 2001US-026961P. 20-MAR-2001; 2001US-0277337P.
XX
XA {CURA-} CURAGEN CORP.

PA (CORT-) COR THERAPEUTICS INC.
XX
XX Burgess CE, Conley PB, Grose WM, Hart M, Kekuda R, Shimkets RA;
P1 Splytek KA, Szekeres ES, Tomlinson JE, Topper JM, Yang R;
XX
XX WPI; 2002-280937/32.
DR P-PSDB; AAU91308.
DR
XX
XX New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
XX Claim 1; Page 98; 263pp; English.

The invention relates to an isolated polypeptide (NOVX) a mature form of NOVX, a NOVX variant (differing by no more than 15%), the nucleotide encoding NOVX (or its complement, fragment or variant), NOVX is NOV1-14, 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it, and antibody against it, are useful for treating or preventing (e.g., by gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy, atherosclerosis, a disorder related to cell signal processing and metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide and nucleic acids are also useful for determining the presence of predisposition to the diseases. The NOVX nucleic acid and polypeptide are especially useful in therapeutic or prophylactic applications for disorders associated with aberrant NOVX expression or activity, e.g. cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus cancer), immune response, graft-versus-host disease, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension, congenital heart defects, multiple sclerosis, inflammation or Albright hereditary osteodystrophy and many other diseases listed in the specification. The DNA encoding the protein is useful in gene therapy for treating the conditions. This is also useful in detection assays, chromosome mapping, tissue typing, diagnostic or prognostic assays, or for developing a powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications. The present sequence encodes a NOVX protein

Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	5, 226-110
Score:	11347.00
Percent Similarity:	99.62%
Best Local Similarity:	99.62%
Query Match:	99.70%
DB:	6
Gaps:	0
Length:	1188
Matches:	260
Conservative:	0
Mismatches:	1
Indels:	0
Indels:	0
Gaps:	0

US-10-043-649-2 (1-261) x ABK61465 (1-1183)

QY	1	Me(GlySerLeu)ProSerArgArgIySerLeuProSerProSerLeuSerSerVal	20
Db	398	ATGGGAAGCTGCTCCAGCAGAAAGAAATCTCTGCCAAGCCCAACCTTGAGTTCCTCTGC	457
QY	21	GIingIyGIingIyProValThiMetGluNagIuaYSerIySalatThraIaValaLeu	40
Db	458	CAAGGCCAGAGGACTGTGTACCATGGAAACAGAGAAACAGAGCCACAGCGGTGGCCCTG	51.7
QY	41	GIYSerPheProAlaGIyGIyProAlaIuLeuSerLeuArgLeuGIyIuProLeuThr	60
Db	518	GGCAGTTTCCCGCAGGTGGCGGCCGAGCTGTGCTGCAACCTCGGGAGCACTTATACC	577
QY	61	IleValSerGIuAspGIyAspTTPTrThrValIeuSerGIuValaSerGIyArgIuTyr	80
Db	578	ATGCTCTGAGAGATGGAGACTGGTGGACGGCTGCTGTAAGCTCTCAGGCAGAGATAT	637
QY	81	AsnIleProSerValHisValAlaIySalSerThiGIyTyrPleuTyrArgIyIuLeuSer	100
Db	638	AACATCCCAAGCGCTCAGCTGGGCAAAATCTTCCATGGGATGAGAGGCGTTAGC	697
QY	101	ArgGIuIySalGIuGIuLeuLeuLeuLeuProGIyAsnProGIyGIyAlaIaPheLeuIle	120
Db	698	AGGAGAAAGCAGAGAACTGCTGTGTGTATCTGGAAACCTCGAGAGGGCGCTTCTCATC	757

QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 758 CGGAGAGCCAGACAGAGAGAGGCTTCTTCTCTGTAAGTCCGCTTACGCCCTCCGCA 817
QY 141 SerTTPAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
DB 818 TCCTGGAGCCGGATGACACACTACAGGATCCACTGCTTACATGCTGCTGCTGATC 877
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180
DB 878 TCACCCGCTTACCTTCCCTTCTTCAAGCTGAGCCCTGGTGAACATTACTTGAGCTGGCG 937
QY 181 AspAspIleCysCysLeuLeuGlyProCysValLeuGlnArgIleGlyProLeuPro 200
DB 938 GATGACATCTGCTGCTTCTTCTTCAAGGAGGCTGTCTCTGAGAGGCTGCTGCTGCTGCT 957
QY 201 GlyValAspIleProLeuProValThrValGlnArgThrProLeuSerTyrPlyGlnLeu 220
DB 998 GGCAGAGATATACCCCTTCTTCTTCAAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAG 1057
QY 221 AspSerSerLeuLeuPheSerGlnAlaAlaThrGlyGlnGlnSerLeuSerGlnGly 240
DB 1058 GACAGCTCCCTCTTCTTCTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1117
QY 241 LeuArgGlnSerLeuSerPheTyrIleSerLeuAspAsnGlyAlaValSerLeuAspAsp 260
DB 1118 CTCGGGAGTCCCTTCTTCTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1177
QY 261 Ala 261
DB 1178 GCC 1180

RESULT 6

ADP82458 standard; DNA; 2788 BP.

ID ADF82458;

ADP82458; (first entry)

26-FEB-2004

Leukemia-related DNA sequence #3014.

Cytostatic; Gene therapy; leukemia; ss.

Unidentified.

NC0203039443-A2.

15-MAY-2003.

04-NOV-2002; 2002WO-EP012303.

05-NOV-2001; 2001EP-00126244.

30-APR-2002; 2002EP-00009758.

(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

(UWU-) UNIV LUDWIG MAXIMILIANS.

(HAFEL) HAFERLACH T.

(SCHO) SCHOCH C.

(KERN) KERN W.

Hafnerlach T, Schoch C, Kern W, Kohlmann A, Schnitger S, Dugas M;

Elis R, Brors B, Mergenthaler S;

WPL; 2003-505037/47.

Determning the subtype of leukemia cells and whether a patient sample

contains leukemia cells or other cells, useful for treating leukemia,

comprises determining the expression profile of a group of markers in a

patient sample.

Disclosure; SEQ ID NO 3014; 2938bp; English.

CC The present invention relates to a method (M1) for determining the
CC subtype of leukemia cells and whether a patient sample contains
CC leukemia cells. The method comprises determining the expression profile
CC of a group of markers in a patient sample. The method is useful for
CC determining the presence of leukemia cells, its types or subtypes, and
CC for the preparation of a medicament for treating leukemia.
SQ Sequence 2788 BP; 663 A; 764 C; 700 G; 595 T; 0 U; 66 Other;

Alignment Scores:

Pred. No.: 2,39e-109

Score: 1345.00

Percent Similarity: 99.62%

Best Local Similarity: 99.62%

Query Match: 99.56%

DB: 10

Gaps: 0

US-10-043-649-2 (1-261) x ADP82458 (1-2788)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20

DB 387 ATGGAGAGTCTGCCAGAT 446

QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerValThrAlaValAlaLeu 40

DB 447 CAAGGCCAG 506

QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60

DB 507 GGCAGTTTCCCGGAG 566

QY 61 IleValSerGlnAspGlyAspTyrPheThrValLeuSerGlnValSerGlyArgGlnTyr 80

DB 567 ATCGTCTGTAGATGAT 626

QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGlnGlyLeuSer 100

DB 627 AACATCCCGAG 686

QY 101 ArgGlnGlyAlaGlnGlnLeuLeuLeuLeuProGlyAsnProGlyAlaIleLeuIle 120

DB 687 AGGAG 746

QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140

DB 747 CGGAG 806

QY 141 SerTTPAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160

DB 807 TCCTGGAGCCGGATGACACACTACAGGATCCACTGCTTGAACATGCTGCTGCTGATC 866

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180

DB 867 TCACCGGCTTCACTTCCCTTCTTCAAGCTGAGCCCTGTGTGAGCATTACTGAGCTGGCG 926

QY 181 AspAspIleCysCysLeuLeuGlyProCysValLeuGlnArgIleGlyProLeuPro 200

DB 927 GATGACATCTGCTGCTTCTTCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 986

QY 201 GlyValAspIleProLeuProValThrValGlnArgThrProLeuSerTyrPlyGlnLeu 220

DB 987 GGCAGAGATATACCCCTTCTTCTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046

QY 221 AspSerSerLeuLeuPheSerGlnAlaAlaThrGlyGlnGlnSerLeuSerGlnGly 240

DB 1047 GACAGCTCCCTCTTCTTCTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1106

QY 241 LeuArgGlnSerLeuSerPheTyrIleSerLeuAspAsnGlyAlaValSerLeuAspAsp 260

DB 1107 CTCGGGAGTCCCTTCTTCTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1166

QY 261 Ala 261

Db 1167 GCC 1169

RESULT 7

AACT7202

AACT7202 standard; cDNA; 837 BP.

AC AACT7202;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.

XX

Human; open reading frame; ORFX; detection; cytosolic; hepatotropic; vlnery; antiproliferative; antiparkinsonian; neurotrophic; anticonvulsant; osteoporotic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antineoplastic; antitumor; antiviral; antibacterial; antifungal; antipneumatic; antihypertensive; antianemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

XX

OS Homo sapiens.

XX

PN W0200058473-A2.

XX

PD 05-OCT-2000.

XX

PF 31-MAR-2000; 2000WO-US008621.

XX

PR 31-MAR-1999; 99US-0127607P.

XX

PR 02-APR-1999; 99US-0127636P.

XX

PR 05-APR-1999; 99US-0127728P.

XX

PR 30-MAR-2000; 2000US-00540763.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach M;

XX

XX WPI, 2000-602362/57.

XX

DR P-PSDB; AAB42993.

XX

PT Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

XX

PS Claim 5; Page 4692-4693; 5507pp; English.

XX

AACT4446 to AACT77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytosolic; hepatotropic; vlnery; antiparkinsonian; neurotrophic; neuroprotective; osteoporotic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antineoplastic; antitumor; antiviral; antifungal; antipneumatic; antihypertensive; antianemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX

XX Sequence 837 BP; 176 A; 254 C; 245 G; 160 T; 0 U; 2 Other;

SQ

Alignment Scores:

Pred. No.: 1,24e-103 Length: 837

Score: 1273.00 Matches: 245

Percent Similarity: 99.60% Conservative: 1

Best Local Similarity: 99.19% Mismatches: 0

Query Match: 94.23% Indels: 0

DB: 3 Gaps: 0

US-10-043-649-2 (1-261) x AACT7202 (1-837)

QY 15 SerLeuSerSerSerValAlaGlnGlyProValThrMetGluAlaGluArgSerLys 34

Db 4 AGCTTGATTTCTCTGCTCCAGGCGCCAGGAGCTTGACCACTGAGAGAGAGAGAG 63

QY 35 AlArhAlaValAlaLeuGlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArg 54

Db 64 GCCACAGCCGCGCCCTGGCAGATTCCCGGAGGTGGCCGCGGAGCTGTCGCTGAGA 123

QY 55 LeuGlyGluProLeuThrIleValSerGluAspGlyAspTyrPheThrValLeuSerGlu 74

Db 124 CTGGGAGCCATTGACCATCTCTGAGATGAGACTGAGTGGACGCTGCTGTGAA 183

QY 75 ValSerGlyArgGluTyrAsnIleProSerValHisValAlaValSerHisGlyTyrP 94

Db 184 GTCTCAGGACAGAGTATATACATCCAGGCTCCACGTCGCAAGTCTCCATGGGTGG 243

QY 95 LeuTyrGluGlyLeuSerArgGluLysAlaGluGluLeuLeuProGlyAsnPro 114

Db 244 CTGTATGAGGCGCTGAG 303

QY 115 GlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerLeuSerVal 134

Db 304 GAGAGGCGCTTCCTCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363

QY 135 ArgLeuSerArgProAlaSerTyrAspArgIleArgHisTyrArgIleHisCysLeuAsp 154

Db 364 CGCCTCAGCGCCGCTGATCTGAGGAGCCGAGTACAGACACTACAGATCACTGCTTAC 423

QY 155 AsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValAsp 174

Db 424 AATGCTGCTGCTGATCATCTCACCGGCTCCTCCTCATCCAGGCGCTGGTGAC 483

QY 175 HisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGluProCysValLeuGln 194

Db 484 CATTACTTGATGCTGGGAGATGACATCTGCTGCTACTCAAGAGAGAGAGAGAGAGAG 543

QY 195 ArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgThrPro 214

Db 544 AGGCTGCGCCGCTCCCTGGAGAGATATACCTTACTGCTGATGCTGACAGAGACCA 603

QY 215 LeuAsnTyrLysGluLeuAspSerSerLeuLeuPheSerGluAlaIleThrGlyGln 234

Db 604 CTCACCTGAGAAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663

QY 235 SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAspGlu 254

Db 664 TCTTTTCTCATGAGAGGCTTCGAGAGTCCCTCAGCTTCAATCAATCAATCAATCAAT 723

QY 255 AlaValSerLeuAspAspAla 261

Db 724 GCTGCTCTTGTGATGATGCC 744

RESULT 8

ABQ9374

ID ABQ9374 standard; cDNA; 1413 BP.

XX

XX ABQ9374;

XX

DT 25-FEB-2003 (first entry)

XX

XX Key Location/Qualifiers
 FT 1..633
 FT /tag= a
 FT /product= "Mouse MARS short isoform protein"
 XX
 XX W0200242452-A2.
 XX
 XX 30-MAY-2002.
 XX
 XX 26-NOV-2001; 2001WO-CA001662.
 XX
 XX 27-NOV-2000; 2000CA-02324663.
 XX
 XX (HOSP-) HOSPITAL FOR SICK CHILDREN.
 XX
 XX Mcglade JC, Loreto MP;
 XX WPT; 2002-566564/60.
 XX P-PSDB; AA015458.
 XX
 XX New isolated modulator of antigen receptor signaling protein or its
 XX fragment, useful for treating malignant disorders such as myeloid
 XX malignancies, autoimmune disorders and myeloproliferative disorders.
 XX
 XX Claim 9; Page 77; 110pp; English.
 XX
 XX The invention comprises the amino acid and coding sequences of modulator
 XX of antigen receptor signaling (MARS) proteins. The MARS protein is a
 XX putative tumor suppressor gene and exhibits structural and sequence
 XX similarity to the Src-like adaptor protein (SLAP). The MARS DNA and
 XX protein sequences of the invention are useful for the treatment of
 XX myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 XX disorders, immunosuppression, myeloproliferative disorders and
 XX malignancies related to the de-regulation of tyrosine kinases (e.g.
 XX breast cancer). The present cDNA sequence encodes a mouse MARS protein
 XX
 XX Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 2,876-97 Length: 737
 XX Score: 1200.50 Matches: 241
 XX Percent Similarity: 92.40% Conservative: 2
 XX Best Local Similarity: 91.63% Mismatches: 1
 XX Query Match: 88.86% Indels: 19
 XX DB: 6 Gaps: 1
 XX
 XX US-10-043-649-2 (1-261) x AAL44090 (1-737)
 QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
 Db 1 ATGGAGACTCTGCCAGAGAAAGAAATCTGCCAACCCCAAGCTTAGTTCTCTGTC 60
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerValThrAlaValAlaLeu 40
 Db 61 CAAAGCCAGGAGCCTGTGACCAATGAGACAGAGAAACCAAGCCACGCTGACCTG 120
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 Db 121 GGCAAGTTCCCGGAGAGTGGCCCGCCGAGCTGTGCTGAGACTCGGGAGGACATTGACC 180
 QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGluTyr 80
 Db 181 ATGCTCTCTGAGAGAGACTGCTGGACGCTGTCTGAAGTCTTACAGCAGAGATAT 240
 QY 81 AsnIleProSerValHisValAlaIleValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
 Db 241 AACATCCCAAGCTCCAGCTGACCAAGCTTCCATGAGTGGCTTATAGAGGAGCTGAGC 300
 QY 101 ArgGluValAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 Db 301 AGGGAAGAAGCAGAGAACTGCTGTGTACTCTGGAAACCCCTGGAGGGGCTTCTCATTC 360

QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 361 CGGAGAGCCAGACAGAGAGGCTTACTCTGTACATGCCCTCAGCCGCTGCA 420
 QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 Db 421 TCTGGAGCCGGAATCAGACACTACAGATCCATGCTTGAACATGCTGTGATCATC 480
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 Db 481 TCACCGGSCCTCACTTCCCTCACTCCAGGCCCTGGTGCACATTA----- 528
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
 Db 529 -----TCTGAGGCTGAGCCGCTCCC 549
 QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
 Db 550 TGGCAAGATATATACCCCTTACTGTGACTGTGCGAGGACACCACTCAACTGGAAGAGCT 609
 QY 220 UAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluLeuSerLeuSerGlyGlu 240
 Db 610 GGACAGCTCCCTCTCTGTTTCTGAAGCTGCCACAGGGAGAGAGTCTTCTCAGTGAAGG 669
 QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAspA 260
 Db 670 TCTCCGGAGATCCCTCAGCTTCTATCATCAGCGCTGAATGAGCAGAGGCTGTCTTGTGATG 729
 QY 260 SPAla 261
 Db 730 ATGCC 734
 XX
 XX RESULT 10
 XX AAS74750
 XX ID AAS74750 standard; cDNA; 2049 BP.
 XX
 XX AAS74750;
 XX
 XX 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #10554.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 XX Homo sapiens.
 XX
 XX W0200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSB-) HYSBQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX WPT; 2001-639362/73.
 XX P-PSDB; ABG10563.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity.
 XX
 XX Claim 1; SEQ ID NO 10554; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 XX sequences. (I) is useful as hybridisation probes, polymerase chain
 XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 1.18e-90 Length: 2049
 Score: 1132.50 Matches: 239
 Percent Similarity: 89.89% Conservative: 1
 Best Local Similarity: 89.51% Mismatches: 4
 Query Match: 83.83% Indels: 23
 DB: 5 Gaps: 1

US-10-043-649-2 (1-261) x AAS74750 (1-2049)

QY 1 MetGlySerLeuProSerArgArgLeuSerLeuProSerLeuSerSerVal 20
 DB ATGGAGAGTCTGCCAGAGAGAAATCTCTGCAAGCCCAAGCTTGAGTCTCTGTC 1024
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 DB CAAAGCCAGGAGCCTGAGCAGTGAAGCAAGAGCAAGGCAAGCCGTGGCCCTG 1084
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 DB GGCACTTCCCGAGAGTGGCCCGCCGAGCTGCTGAGACCTCGGAGGCCATTGACC 1144
 QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
 DB ATCTCTCTGAGTGAAGTGAAGTGTGACGCTGCTGTGAAGTCTCAGGAGAGAGTAT 1204
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrPheLysGlyLeuSer 100
 DB AACATCCCAAGCCGTCACGTCGCAAGATCCCATGGTGGCTGTATGAGGGCCTGAGC 1264
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 DB AGGAGAAAGAGAGAACTGCTGTTGTTACTGGAAACCTGAGAGGGGCTTCTCATC 1324
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 DB CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCAGTCCCGCTCAACCGCCCTGCA 1384
 QY 141 SerTyrPheArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheLysIle 160
 DB TTTCGGAGCCGAGACAGACACTACAGATTCACGCTTGACATGGCTGGCTGTACATC 1444
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 DB TCACCGGCTCCTACCTTCCCTCTCATCCAGGCTGAGGACCACTTAC----- 1492
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuPr 200
 DB 1493 -----TCTGAGGGCTGGCCGCTGCC 1513
 QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrPheGluLe 220

DB 1514 TGCGAAGATATATACCTTACTGACTGTGACAGAGACACCACTCACTGAAAGAGCT 1573
 QY 220 uAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuLeu-SerGlu 239
 DB 1574 GGAACGTCCTCTCTGTTTCTGAAAGTGCACAGGGGAGAGATCTTCTTCAAGAGAG 1633
 QY 240 GlyLeuArgGluSer-LeuSerPheTyrIleSer-LeuAsnAspGluAlaValSer-Leu 258
 DB GGGTCCGGAGATCCCTCAGCTTCTTACATCAGCCCTGAATAGACAGAGCTGTCTTTTG 1693
 QY 259 AspAspAla 261
 DB 1694 GATGATGCC 1702

RESULT 11

AA144087
 ID AA144087 standard; cDNA; 1348 BP.

AC AA144087;
 XX 03-OCT-2002 (first entry)

DE Mouse modulator of antigen receptor signalling protein coding sequence.

XX Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
 KW MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;
 KW myeloid malignancy; acute myelogenous leukemia; autoimmune disorder;
 KW immunosuppression; myeloproliferative disorder; breast cancer.

XX Mus sp.

XX Key Location/Qualifiers
 FH CDS
 FT 282..1061
 FT /*tag= a
 FT /product= "Mouse MARS protein"

XX W0200242452-A2.

XX PD 30-MAY-2002.

XX PF 26-NOV-2001; 2001WO-CM001662.

XX PR 27-NOV-2000; 2000CA-02324663.

XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX PI Meglade JC, Loreto MP;

XX DR WPI; 2002-566564/60.

XX DR P-PSDB; AAO15456.

XX PT New isolated modulator of antigen receptor signaling protein or its
 PT fragment, useful for treating malignant disorders such as myeloid
 PT malignancies, autoimmune disorders and myeloproliferative disorders.

XX PS Claim 10; Fig 1A; 110bp; English.

XX The invention comprises the amino acid and coding sequences of modulator
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
 CC putative tumour suppressor gene and exhibits structural and sequence
 CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
 CC protein sequences of the invention are useful for the treatment of
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 CC disorders, immunosuppression, myeloproliferative disorders and
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.
 CC breast cancer). The present cDNA sequence encodes a mouse MARS protein

XX Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.73e-82 Length: 1348
 Score: 1032.00 Matches: 209
 Percent Similarity: 85.88% Conservative: 16

Best Local Similarity: 79.77%
Query Match: 76.39%
Db: 6
Gaps: 3
US-10-043-649-2 (1-261) x AAL4087 (1-1348)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 282 ATGGGAAGTTGTCCAGACAGAGGAAAC--TCCACCCCCAGCCCACTCTCTGCT 338
QY 21 GlnGlyGlnGlyProValTherMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db 339 CCAAGCCAGGAACCCCTGTCTTCCATGACCAAGAAACACACAGTCAAGCTGTGCTG 398
QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60
Db 399 GGCAGTTTCCAGCAGGTGGAACAGGCCAGACTATCTCGAAGCTCGGGAGCCCTGACC 458
QY 61 IleValSerGlnAspGlyAspTyrTrpThrValLeuSerGlnValSerGlyArgGlnTyr 80
Db 459 ATCATCTCTGAGAGGAGATGAGTGTGGAACAGTCCAGTCCGAACTCTCAGGACAGAGTAC 518
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrPheLeuTyrGlnGlyLeuSer 100
Db 519 CACATGCCCACTGTGTATGTGCTTAAGTCCGCCAGGAGTGTGTACAGAGGCTGAGC 578
QY 101 ArgGlnValAlaGlnGlnLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 579 CGGGAAGAGCCGAGGAACACTCTCTGTACCTGGGAACCCGAGAGGGCTCTCTATC 638
QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 639 CGGAGAGCCAGACAGAGAGAGGCTGTATTCCTCTCCGCTCCGACTACAGCCCTGCA 698
QY 141 SerTyrAspArgGlyLeuArgHisTyrArgTyrLeuHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 699 TCTTGGAGCCGAGTACAGACTACAGATACAGCTCTTGAACATGGCTGTGATACATC 758
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180
Db 759 TCACCTGGCTCACCTTCCCTCACTCCAGGCTTGGTGGAGCATTAATCTGAGGTAGCA 818
QY 181 AspAspIleCysCysLeuLeuGlnProCysValLeuGlnArgAlaGlyProLeuPro 200
Db 819 GATGACATCTGCTGCTCCCTCAGGAGGCTGTGTCTCGAGAAAGCTTGGCCCACTACT 878
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAspTyrLeuGln 220
Db 879 GGCAGAGATACCTCTCCACCTGTGCTGTGCCAACATCATCACTPAATTGGAAAAAGCTG 938
QY 221 AspSerSerLeuLeuPheSerGlnAla--AlaThrGlyGlnGlnSerLeuLeuSerGln 239
Db 939 GACCGCAGACCTCTCTTCTGGAAGCACCTCGAGATGGGAGGACATCTCTCTACGTAG 998
QY 240 GlyLeuArgGlnSerLeuSerPheTyrIleSerLeuAsnAspGlnAlaValSerLeuAsp 259
Db 999 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCGTGCTGAGGAC-----CCCTTGAT 1052
QY 260 AspAla 261
Db 1053 GATGCT 1058

RESULT 12
ABQ98670
ID ABQ98670 standard; DNA; 763 BP.
XX
XX ABQ98670;
XX
XX 04-NOV-2002 (first entry)
XX
XX Human ORF47 coding sequence.
XX
XX Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;

KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
OS Homo sapiens.
PN US2002082206-A1.
XX
XX 27-JUN-2002.
XX
XX 30-MAY-2001; 2001US-00867550.
XX
XX 30-MAY-2000; 2000US-0208427P.
XX
XX (LEACH/) LEACH M D.
XX (MEHR/) MEHRABAN F.
XX (CONL/) CONLEY P B.
XX (TOPP/) TOPPER J N.
XX (LAWD/) LAW D.
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
XX WPI; 2002-626554/67.
DR P-PSDB; ABP64107.
XX
XX New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
XX Claim 2; SEQ ID NO 953; 78bp; English.
XX
XX The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer, blood
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?docID=2002082206
XX
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,36e-64 Length: 763
Score: 830.00 Matches: 159
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.44% Indels: 0
Gaps: 0

US-10-043-649-2 (1-261) x ABQ98670 (1-763)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20
Db 286 ATGGGAAGTTGTCCAGACAGAGGAAAC--TCTGCCAGGCCCAAGTTCCTCTGTC 345
QY 21 GlnGlyGlnGlyProValTherMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db 346 CAAAGCCAGGAACCTGTGACCATGACAGCAAGAAACAGGCCCAAGCCGCTGCTG 405
QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60
Db 406 GGCAGTTTCCAGCAGGTGAGCCCGGAGCTGTGCTGAGACTCGGGAGACCATTTGACC 465
QY 61 IleValSerGlnAspGlyAspTyrTrpThrValLeuSerGlnValSerGlyArgGlnTyr 80

Db 466 ATGCTCTGAGATGAGACCTGTGAGCGGTGCTGTAAGTCTCAGGACAGAGATAT 525
QY 81 AsnIleProSerValHisValAlaIalysValSerHisGlyTyrPleuTyrGluIleuSer 100
Db 526 AACATCCCCAGCGCTCCACCGTGGCCAAAGTCCCATGGGTGCTGTATGAGGGCCCTGAGC 585
QY 101 ArgGluIysAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
Db 586 AGGAGAAAGACAGAGAACTGCTGTTCTTACTCTGAGAACCTTGAGGGGCGCTTCCTCATC 645
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerIleuSerValArgIleuSerArgProAla 140
Db 646 CGGAGAGCCAGACAGAGAAAGGCTTACTCTCTGTCAGCTCCGCTCCAGCCGCCCTGCA 705
QY 141 SerThrAspArgGlyLeuArgHisTyrArgIleHisCysIleuAspAsnGlyTyrPleuTyr 159
Db 706 TCTCGGAGCCGATCAGACACTACAGATCCACTGCTTGAACATGGCTGCTGTAC 762
RESULT 13
ID AAS74748 standard; cDNA; 603 BP.
XX AAS74748;
AC AAS74748;
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #10552.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX PF 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSB-) HYSBQ INC.
XX Demanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
DR P-PSDB; ABG10561.
XX
XX New isolated polynucleotide and encoded polypeptides; useful in
PT diagnostics; forensics; gene mapping; identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 10552; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS6197-AAS94564 represent novel human diagnostic

CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences
XX
XX Sequence 603 BP; 124 A; 189 C; 164 G; 126 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1,07e-55 Length: 603
Score: 731.00 Matches: 151
Percent Similarity: 82.45% Conservative: 4
Best Local Similarity: 80.32% Mismatches: 21
Query Match: 54.11% Indels: 12
DB: Gaps: 4
US-10-043-649-2 (1-261) x AAS74748 (1-603)
QY 78 ArgGluTyrAsnIleProSerValHisValAlaIalysValSerHisGlyTyrPleuTyrGlu 97
Db 61 CGCAGGAGGAGAACTGGCCCAAGAGATCGGTG---AAGATCCCTCCGGGACTGGCTGTAC--- 114
QY 98 GlyIleuSerArgGluIysAla-----GluGluLeuLeuLeuProGlyAsn 113
Db 115 ---TTGACCCGCTAACACGCTTACCCCTCAGAGCAGAGAGCTGAGCTTTCTGGACAG 171
QY 114 ProGlyGlyAlaPheLeuIleArgGluSerGlnThrArgArgGlySerTyrSerIleuSer 133
Db 172 ACCAACCTGTCAAGTCTG-----CAAGACATGGCTCTTACTCTCTGTCA 216
QY 134 ValArgIleuSerArgProAlaSerTyrPaspArgIleArgHisTyrArgIleHisCysIleu 153
Db 217 GTCCGCTCAGCCGCTGCTGATCTGGGACCGGATCGACATCAACAGATCCACTGCTT 276
QY 154 AspAsnGlyTyrPleuTyrIleSerProArgLeuThrPheProSerIleuGlnAlaIleuVal 173
Db 277 GACATAGCTGCGCTGTACATCTACCGGCTTACCTTCCCTTCACTCCAGGCTTGTG 336
QY 174 AspHisTyrSerGluLeuAlaAspAspIleCysCysIleuLeuIysGluProCysValIleu 193
Db 337 GACATTACTCTGAGCTGAGCGGATGACATCTGCTGCTTCAAGAGCCCTGTGCTCTG 396
QY 194 GlnArgAlaGlyProIleuProGlyIysAspIleProIleuProValThrValGlnArgThr 213
Db 397 CAGAGGCTGCGCCGCTCCCTGGCCAGAGATATACCTTACTGTGACTGTGCAGAGACA 456
QY 214 ProIleuAsnTyrIysGluLeuAspSerSerIleuLeuPheSerGluAlaIleuGlu 233
Db 457 CCACTCAACTGAAAGAGCTGAGACGCTCCCTCTGTTTCTTAAGCTGCCACAGGGAG 516
QY 234 GluSerLeuLeuSerGluGlyLeuArgGluSerIleuSerPheTyrIleSerIleuAsnAsp 253
Db 517 GAGCTCTTCTCAAGTAGGGGTCTCCGGAGTCCCTCAGCTTCTTACATCAGGCTGAATGAC 576
QY 254 GluAlaValSerIleuAspAspAla 261
Db 577 GAGGCTGTCTCTTGTGATGATGCC 600
RESULT 14
ID ADL63090 standard; DNA; 864 BP.
XX ADL63090;
AC ADL63090;
DT 20-MAY-2004 (first entry)
XX
XX Human ovarian cancer DNA marker #21302.
XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX Homo sapiens.
XX
XX WO200170979-A2.
XX

PD	27-SEP-2001.
XX	
PF	21-MAR-2001; 2001MO-US009126.
XX	
PR	21-MAR-2000; 2000US-0191031P.
PR	25-MAY-2000; 2000US-0207124P.
PR	15-JUN-2000; 2000US-0211940P.
PR	07-JUL-2000; 2000US-0216820P.
PR	25-JUL-2000; 2000US-0220661P.
PR	21-DEC-2000; 2000US-0257672P.
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Lee J, Lillie J;
XX	
XX	WPI; 2001-611502/70.
DR	
XX	
PT	Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT	cancer cells as compared to their normal non-cancerous ovarian cells are
PT	used to characterize stage, grade, histological type of ovarian cancer.
XX	
ES	Disclosure; SEQ ID NO 21302; 106pp; English.

Alignment Scores:	
Pred. No.:	4.9e-48
Score:	647.00
Percent Similarity:	99.22%
Best Local Similarity:	98.45%
Query Match:	47.89%
EB:	5
Length:	864
Matches:	127
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-10-043-649-2 (1-261) X ADL63090 (1-864)

QY	1	MetGlySerLeuProSerThrArgTrpGlySerLeuProSerProSerLeuSerSerVal	20
Db	450	ATGGAGAGCTCTCCACGAGAAAGAACTCTGGCCAAACCCAAAGTTGAGTTCCTGTGC	509
QY	21	GlnGlyGlnGlyProValThrMetCysLysIleGluArgSerLysAlaThrAlaValAlaLeu	40

Db	510	CAAGGCCAGGGACCTGTGACCATGGAGCAGAGAGAAGCAAGGCCACAGCCGTGGCCCTG	563
Qy	41	GLYSERPhEPProAlaGLYGLYProAlaGLULEuSERLeuArgLEuGLYGLYUProLeuThr	60
Db	570	GGCATATTTCCGGCAGGTGGCCCGGCCACAGCTGTGGCTAGACTCGGGAGGCCATTGGACC	629
Qy	61	ILEuLSErGLuASPGLYASPrTrpTrpThrValILEuSErGLYValSErGLYArGGLUTyr	80
Db	630	ATCGCTCTGAGGATGGAGACTGTGGAGCGGTGCTCTGTGAAGTCTCAGGCAGAGAGATAT	685
Qy	81	ASnLIEProSErValHISValAlaLysValSErHisGLYTrPLEuTYrGLUGLYLeuSEr	100
Db	690	AACATCCCCAGGCTCCAGCTGCCCAAAGCTCCCATGGGTGGCTGTATATGGGGCCCTGAGC	749
Qy	101	ArgGLYsAlaGLUGLUleuDeuLeuLeuProGLYASnProGLYGLYAlaPheLeuILE	120
Db	750	AGGGAGAAAGCAGAGAACTGTGTGTGTACTGTGGAACTCTGGAGGGGCGCTTCTCATTC	809
Qy	121	ArgGLuSErGLInThrArgArgGLYSEr	129
Db	810	CGGAGAGCCACAGCCAGGAAAGAGGCTC	836

RESULT 15

ID ABQ99151 standard; DNA; 875 BP.

AC ABQ99151;

DT 04-NOV-2002 (first entry)

Human ORF958 coding sequence.

KM Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KM Antiinflammatory; gene therapy; human; OREX; atberogenic; platelet;
KM human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque
KM cancer; cardiovascular disease; allergy; autoimmune disease;
KM wound healing; blood coagulation disorder; inflammatory disorder; ds

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PF 30-MAY-2001; 2001US-00867550.

PR 30-MAY-2000; 2000US-0208427P.

PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPE/) TOPPER J N.
PA (LAWD/) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D,

DR WPI; 2002-626554/67.

XXXXXX

PT New polypeptide designated ORRX are present in human atherogenic cells
PT and are useful to prevent and treat ORRX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.

PS Claim 2; SEQ ID NO 1915; 78pp; English.

CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (AB963631-AB964681 and AB968194-AB969267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological

CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?docID=2002082206

XX
SQ Sequence 875 BP; 205 A; 259 C; 225 G; 185 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.3e-42	Length:	875
Score:	586.00	Matches:	112
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	43.38%	Indels:	0
DB:	6	Gaps:	0

US-10-043-649-2 (1-261) x ABQ99151 (1-875)

QY	150	IleHisCysIeuAspAsnGlyTyrLeuTyrIleSerProArgLeuThrPheProSerLeu	169
DB	4	ATCCACTGCTTGAACATGGCTGGCTGTACATCTCAACCGGCTCACTTCCCTCACTC	63
QY	170	GlnAlaIeuValAspHisTyrSerGluLeuAlaAspAspIleCysCysLeuLeuLysGlu	189
DB	64	CAGGCCCTGGTGGACCATCTACTGAGCTGGCGGATGACATCTGCTGCTACTCAAGAG	123
QY	190	ProCysValIleuGlnArgAlaGlyProLeuProGlyLysAspIleProLeuProValThr	209
DB	124	CCTGTGTCTTGGAGAGGGCTGGCCGCTCCCTGGGCAAGGATATACCCCTTACCTGTGACT	183
QY	210	ValGlnArgThrProLeuAsnTyrPlyGluLeuAspSerSerLeuLeuPheSerGluAla	229
DB	184	GTGCAGAGGACACCACTCAACTGGAAGAAGCTGGACAGCTCCCTCTGTTTCTGAAGCT	243
QY	230	AlaThrGlyGluGlnSerLeuLeuSerGluGlyLeuArgGlnSerLeuSerPheTyrIle	249
DB	244	GCCACAGGGGAGGAGTCTCTTCTCAGTGAAGGATCTCCGGGAGTCCCTCAGCTTCTACATC	303
QY	250	SerLeuAsnAspGlnAlaValSerLeuAspAspAla	261
DB	304	AGCTGATATGACGAGGCTGTCTCTTTGGATGATGCC	339

Search completed: November 17, 2004, 00:59:18
Job time : 539 secs

CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
 CC putative tumour suppressor gene and exhibits structural and sequence
 CC similarity to the Src-like adaptor protein (SLAP). The MARS DNA and
 CC protein sequences of the invention are useful for the treatment of
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 CC disorder, immunosuppression, myeloproliferative disorders and
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.
 CC breast cancer). The present cDNA sequence encodes a human MARS protein
 XX

XX Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Query Match 100.0%; Score 786; DB 6; Length 786;
 Best Local Similarity 100.0%; Pred. No. 1.1e-197;

Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGAAGTCTGCCACAGAAAGAAATCTTGTCCAGACCCCAAGCTTGAAGTCTCTCTGTC 60
DB 1 ATGGGAAGTCTGCCACAGAAAGAAATCTTGTCCAGACCCCAAGCTTGAAGTCTCTCTGTC 60
QY 61 CAAGGCCAGGGACCTGTGACCATGTAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 CAAGGCCAGGGACCTGTGACCATGTAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 GGCAGATTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTGGGGAGAGCAATTGACC 180
DB 121 GGCAGATTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTGGGGAGAGCAATTGACC 180
QY 181 ATGCTCTCTAGATGAGTGAAGTGTGTGAGCGGTGTGTCTGAAGCTTGAAGCTGAGAGATAT 240
DB 181 ATGCTCTCTAGATGAGTGAAGTGTGTGAGCGGTGTGTCTGAAGCTTGAAGCTGAGAGATAT 240
QY 241 AACATCCCGACGCTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGGCTGAGC 300
DB 241 AACATCCCGACGCTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGGCTGAGC 300
QY 301 AGGAGAGAAAGCAGAGAGAGTGTGTGTACTGTGTGAGAAACCTGGAGGGGCTTCTCATC 360
DB 301 AGGAGAGAAAGCAGAGAGAGTGTGTGTACTGTGTGAGAAACCTGGAGGGGCTTCTCATC 360
QY 361 CGGAGAGCCAGACCCAGAGAGAGGCTTACTCTGTGACGCGCCCTCAGCGGCGCTGCA 420
DB 361 CGGAGAGCCAGACCCAGAGAGAGGCTTACTCTGTGACGCGCCCTCAGCGGCGCTGCA 420
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGGCTGTATCATC 480
DB 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGAACAATGGCTGGCTGTATCATC 480
QY 481 TCACCGGCTCTACCTTCCCTCTCACTCCAGGCTTGTGTGACATTTACTCTGAGCTGGCG 540
DB 481 TCACCGGCTCTACCTTCCCTCTCACTCCAGGCTTGTGTGACATTTACTCTGAGCTGGCG 540
QY 541 GATGACATCTGCTGCTCTCACTCAAGAGAGGCTGTGTCTCTGAGAGAGGCTGGCGCTCCCT 600
DB 541 GATGACATCTGCTGCTCTCACTCAAGAGAGGCTGTGTCTCTGAGAGAGGCTGGCGCTCCCT 600
QY 601 GGCAGAGATATACCCCTACCTGTGATGTGAGAGAGACACCACTCAACTGGAAGAGAGCTG 660
DB 601 GGCAGAGATATACCCCTACCTGTGATGTGAGAGAGACACCACTCAACTGGAAGAGAGCTG 660
QY 661 GACAGCTCCCTCTCTGTTTCTGAGAGCTGGCCACAGGGGAGGAGTCTTCTTCAAGAGAGGT 720
DB 661 GACAGCTCCCTCTCTGTTTCTGAGAGCTGGCCACAGGGGAGGAGTCTTCTTCAAGAGAGGT 720
QY 721 CTCGGGAGAGTCCCTCAGCTTCTATCATCAGCTGATGACGAGGGCTCTCTTTGGATGAT 780
DB 721 CTCGGGAGAGTCCCTCAGCTTCTATCATCAGCTGATGACGAGGGCTCTCTTTGGATGAT 780
QY 781 GCCTAG 786
DB 781 GCCTAG 786

```

RESULT 2

AB074343
 ID AB074343 standard; cDNA; 786 BP.
 XX
 AC AB074343;
 XX
 DT 15-OCT-2002 (first entry)
 XX
 DE Human Src-like inhibitory molecule (SLIM) encoding cDNA.
 XX
 KW Human; Src-like inhibitory molecule; SLIM; Src-like adapter protein;
 KW SLAP; inhibitor; antiinflammatory; immunosuppressive; anti-HIV;
 KW modulator; lymphocyte; Cbl; gene therapy; immunodeficiency disorder;
 KW acquired immune deficiency syndrome; AIDS; acute inflammatory disorder;
 KW chronic inflammatory disorder; autoimmune disorder; transplant rejection;
 KW gene; ss.
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 1..786
 FT /*tag= a
 FT /product= "SLIM"
 FT /note= "Src-like inhibitory molecule"
 XX
 PN WO20025707-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 10-JAN-2002; 2002WO-US000718.
 XX
 PR 10-JAN-2001; 2001US-0260953P.
 XX
 PA (RIGI-) RIGEL PHARM INC.
 XX
 PI Holland ST, Mendenhall MK, Pardo J, Spencer C, Fu AC, Luo Y;
 PI Payan DG, Mancebo HS, Wu J, Zhou X, Shen M, Liao XC, Sheng N;
 XX
 DR WPI; 2002-575432/61.
 XX
 DR P-PSDB; ABP52187.
 XX
 PT New src-like inhibitory molecule protein, useful for treating
 PT immunodeficiency disorders and inflammatory disorders, comprises N-
 PT terminal myristylation sequence, SH2 domain and/or SH3 domain.
 XX
 PS Claim 3; Fig 2A; 91pp; English.
 XX
 CC The present sequence encodes the human Src-like inhibitory molecule
 CC (SLIM) protein (I). The present invention describes a SLIM protein
 CC comprising an N-terminal myristylation sequence, an N-terminal SH2
 CC domain, and an N-terminal SH3 domain which can bind to Cbl, or comprising
 CC an N-terminal myristylation sequence and an N-terminal SH2 domain which
 CC is unable to bind to Cbl. (I) has antiinflammatory, immunosuppressive and
 CC anti-HIV activities, and can be used as a modulator of lymphocyte
 CC activation, and of ubiquitination of a Cbl target protein, and in gene
 CC therapy. (I) is useful for screening a bioactive agent capable of binding
 CC to SLIM. (I) is also useful for screening a bioactive agent capable of
 CC modulating SLIM binding. (I) or its fragments is useful in the study or
 CC in the treatment of conditions which involves this function or
 CC dysregulation of SLIM protein activity, i.e. to diagnose, treat or
 CC prevent SLIM associated disorders. (I) or the polynucleotide encoding it
 CC (II) is useful for modulating leukocyte and/or platelet activation, for
 CC modulating antigen receptor-induced signalling and activation in
 CC leukocyte and/or platelets and for modulating antigen receptor-induced
 CC signalling and activation in lymphocytes and/or mast cells. (I) or (II)
 CC is also useful for modulating the basal activity of lymphocytes. (I) or
 CC (II) is useful in the treatment of immunodeficiency disorders, such as
 CC acquired immunodeficiency syndrome (AIDS), for the prevention and
 CC treatment of acute inflammatory disorders, chronic inflammatory
 CC disorders, autoimmune disorder and transplant rejection
 CC
 XX
 SQ Sequence 786 BP; 162 A; 234 C; 231 G; 159 T; 0 U; 0 Other;

Query Match 100.0%; Score 786; DB 6; Length 786;

QY	121	GGCAGTTTCCCGGAGGTGGCCCGGACGCTGCGCTGAGACTCCGGGAGACCAATTACC	180
Db	535	GGCAGTTTCCCGGAGGTGGCCCGGACGCTGCGCTGAGACTCCGGGAGACCAATTACC	594
QY	181	ATCGTCTCTGAGATGAGAGACTGGTGGACGGTGTCTGTTAAGTCTCAGGCAGAGATAT	240
Db	595	ATCGTCTCTGAGATGAGAGACTGGTGGACGGTGTCTGTTAAGTCTCAGGCAGAGATAT	654
QY	241	AACATCCCAAGCCGTCACGTGGCCCAAGTCCCATGGTGGCTGTATGAGAGGCTAGC	300
Db	655	AACATCCCAAGCCGTCACGTGGCCCAAGTCCCATGGTGGCTGTATGAGAGGCTAGC	714
QY	301	AGGAGAAAGCAGAGAGAACTGCTGTGTTTACCTGGAAACCTTGGAGGGGCTTCCATC	360
Db	715	AGGAGAAAGCAGAGAGAACTGCTGTGTTTACCTGGAAACCTTGGAGGGGCTTCCATC	774
QY	361	CGGAGAGCCAGACCCAGAGAGGCTTTACTTCTCTGTAGTTCGCGCTTACGCCCTTGCA	420
Db	775	CGGAGAGCCAGACCCAGAGAGGCTTTACTTCTCTGTAGTTCGCGCTTACGCCCTTGCA	834
QY	421	TCTTGGAGCCGGATCAGACCTACAGATCCACTGCTTGAACAAATGGCTGGCTGTACATC	480
Db	835	TCTTGGAGCCGGATCAGACCTACAGATCCACTGCTTGAACAAATGGCTGGCTGTACATC	894
QY	481	TCACCGGECTCACTTCTCCCTCACTCCAGGCTCTGTGTGACCAATTACTTGAAGTGGCG	540
Db	895	TCACCGGECTCACTTCTCCCTCACTCCAGGCTCTGTGTGACCAATTACTTGAAGTGGCG	954
QY	541	GATGACATCTGCTGCTACTCAAGGAGGCTGTGTCTGCAAGAGGGCTGGCCGCTCCCT	600
Db	955	GATGACATCTGCTGCTACTCAAGGAGGCTGTGTCTGCAAGAGGGCTGGCCGCTCCCT	1014
QY	601	GGCAAGATATACCCCTTACTGTGTGACTGTGCAAGAGACCACTCAACTGAAAGAGCTG	660
Db	1015	GGCAAGATATACCCCTTACTGTGTGACTGTGCAAGAGACCACTCAACTGAAAGAGCTG	1074
QY	661	GACAGCTCCCTCCGTTTCTTGAAGCTGCCACAGGGGAGAGTCTTCTCAGTAGAGAT	720
Db	1075	GACAGCTCCCTCCGTTTCTTGAAGCTGCCACAGGGGAGAGTCTTCTCAGTAGAGAT	1134
QY	721	CTCCGGGAGTCCCTCAGCTTCTTACATCAAGCTGAAATGACAGAGCTGTCTCTTTGGATAT	780
Db	1135	CTCCGGGAGTCCCTCAGCTTCTTACATCAAGCTGAAATGACAGAGCTGTCTCTTTGGATAT	1194
QY	781	GGCTTG 786	
Db	1195	GGCTTG 1200	

Accession	Gene	Protein	Species	Feature	Sequence	Length	GC Content	GC Skew	GC Bias	GC Bias2	GC Bias3	GC Bias4	GC Bias5	GC Bias6	GC Bias7	GC Bias8	GC Bias9	GC Bias10	GC Bias11	GC Bias12	GC Bias13	GC Bias14	GC Bias15	GC Bias16	GC Bias17	GC Bias18	GC Bias19	GC Bias20	GC Bias21	GC Bias22	GC Bias23	GC Bias24	GC Bias25	GC Bias26	GC Bias27	GC Bias28	GC Bias29	GC Bias30	GC Bias31	GC Bias32	GC Bias33	GC Bias34	GC Bias35	GC Bias36	GC Bias37	GC Bias38	GC Bias39	GC Bias40	GC Bias41	GC Bias42	GC Bias43	GC Bias44	GC Bias45	GC Bias46	GC Bias47	GC Bias48	GC Bias49	GC Bias50	GC Bias51	GC Bias52	GC Bias53	GC Bias54	GC Bias55	GC Bias56	GC Bias57	GC Bias58	GC Bias59	GC Bias60	GC Bias61	GC Bias62	GC Bias63	GC Bias64	GC Bias65	GC Bias66	GC Bias67	GC Bias68	GC Bias69	GC Bias70	GC Bias71	GC Bias72	GC Bias73	GC Bias74	GC Bias75	GC Bias76	GC Bias77	GC Bias78	GC Bias79	GC Bias80	GC Bias81	GC Bias82	GC Bias83	GC Bias84	GC Bias85	GC Bias86	GC Bias87	GC Bias88	GC Bias89	GC Bias90	GC Bias91	GC Bias92	GC Bias93	GC Bias94	GC Bias95	GC Bias96	GC Bias97	GC Bias98	GC Bias99	GC Bias100	GC Bias101	GC Bias102	GC Bias103	GC Bias104	GC Bias105	GC Bias106	GC Bias107	GC Bias108	GC Bias109	GC Bias110	GC Bias111	GC Bias112	GC Bias113	GC Bias114	GC Bias115	GC Bias116	GC Bias117	GC Bias118	GC Bias119	GC Bias120	GC Bias121	GC Bias122	GC Bias123	GC Bias124	GC Bias125	GC Bias126	GC Bias127	GC Bias128	GC Bias129	GC Bias130	GC Bias131	GC Bias132	GC Bias133	GC Bias134	GC Bias135	GC Bias136	GC Bias137	GC Bias138	GC Bias139	GC Bias140	GC Bias141	GC Bias142	GC Bias143	GC Bias144	GC Bias145	GC Bias146	GC Bias147	GC Bias148	GC Bias149	GC Bias150	GC Bias151	GC Bias152	GC Bias153	GC Bias154	GC Bias155	GC Bias156	GC Bias157	GC Bias158	GC Bias159	GC Bias160	GC Bias161	GC Bias162	GC Bias163	GC Bias164	GC Bias165	GC Bias166	GC Bias167	GC Bias168	GC Bias169	GC Bias170	GC Bias171	GC Bias172	GC Bias173	GC Bias174	GC Bias175	GC Bias176	GC Bias177	GC Bias178	GC Bias179	GC Bias180	GC Bias181	GC Bias182	GC Bias183	GC Bias184	GC Bias185	GC Bias186	GC Bias187	GC Bias188	GC Bias189	GC Bias190	GC Bias191	GC Bias192	GC Bias193	GC Bias194	GC Bias195	GC Bias196	GC Bias197	GC Bias198	GC Bias199	GC Bias200	GC Bias201	GC Bias202	GC Bias203	GC Bias204	GC Bias205	GC Bias206	GC Bias207	GC Bias208	GC Bias209	GC Bias210	GC Bias211	GC Bias212	GC Bias213	GC Bias214	GC Bias215	GC Bias216	GC Bias217	GC Bias218	GC Bias219	GC Bias220	GC Bias221	GC Bias222	GC Bias223	GC Bias224	GC Bias225	GC Bias226	GC Bias227	GC Bias228	GC Bias229	GC Bias230	GC Bias231	GC Bias232	GC Bias233	GC Bias234	GC Bias235	GC Bias236	GC Bias237	GC Bias238	GC Bias239	GC Bias240	GC Bias241	GC Bias242	GC Bias243	GC Bias244	GC Bias245	GC Bias246	GC Bias247	GC Bias248	GC Bias249	GC Bias250	GC Bias251	GC Bias252	GC Bias253	GC Bias254	GC Bias255	GC Bias256	GC Bias257	GC Bias258	GC Bias259	GC Bias260	GC Bias261	GC Bias262	GC Bias263	GC Bias264	GC Bias265	GC Bias266	GC Bias267	GC Bias268	GC Bias269	GC Bias270	GC Bias271	GC Bias272	GC Bias273	GC Bias274	GC Bias275	GC Bias276	GC Bias277	GC Bias278	GC Bias279	GC Bias280	GC Bias281	GC Bias282	GC Bias283	GC Bias284	GC Bias285	GC Bias286	GC Bias287	GC Bias288	GC Bias289	GC Bias290	GC Bias291	GC Bias292	GC Bias293	GC Bias294	GC Bias295	GC Bias296	GC Bias297	GC Bias298	GC Bias299	GC Bias300	GC Bias301	GC Bias302	GC Bias303	GC Bias304	GC Bias305	GC Bias306	GC Bias307	GC Bias308	GC Bias309	GC Bias310	GC Bias311	GC Bias312	GC Bias313	GC Bias314	GC Bias315	GC Bias316	GC Bias317	GC Bias318	GC Bias319	GC Bias320	GC Bias321	GC Bias322	GC Bias323	GC Bias324	GC Bias325	GC Bias326	GC Bias327	GC Bias328	GC Bias329	GC Bias330	GC Bias331	GC Bias332	GC Bias333	GC Bias334	GC Bias335	GC Bias336	GC Bias337	GC Bias338	GC Bias339	GC Bias340	GC Bias341	GC Bias342	GC Bias343	GC Bias344	GC Bias345	GC Bias346	GC Bias347	GC Bias348	GC Bias349	GC Bias350	GC Bias351	GC Bias352	GC Bias353	GC Bias354	GC Bias355	GC Bias356	GC Bias357	GC Bias358	GC Bias359	GC Bias360	GC Bias361	GC Bias362	GC Bias363	GC Bias364	GC Bias365	GC Bias366	GC Bias367	GC Bias368	GC Bias369	GC Bias370	GC Bias371	GC Bias372	GC Bias373	GC Bias374	GC Bias375
-----------	------	---------	---------	---------	----------	--------	------------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------

XX WP1, 2003-821598/77.
DR
XX Hepatic fibrosis disease markers comprising polynucleotides or
PT antibodies, useful for improved diagnosis, screening and developing drugs
PT to treat hepatitis, to control cirrhosis and carcinoma.
XX
PS Claim 1; SEQ ID NO 203; 313pp; Japanese.
CC The present invention relates to hepatic-fibrosis disease markers
CC (ADPF90539-ADPF90871) and related proteins (ADPF90872-ADPF90917). The
CC sequences are useful for detecting and treating hepatic fibrosis caused
CC by alcohol consumption, virus infection, etc., and the associated chronic
CC hepatitis, etc. leading to liver cirrhosis and hepatic carcinoma. The
CC markers allow the cause of hepatic fibrosis to be clarified (diagnostic
CC precision), so more suitable treatments can be developed and given.
XX
SQ Sequence 3646 BP; 782 A; 954 C; 1031 G; 877 T; 0 U; 2 Other;

	Query Match	100.0%	Score 786;	DB 10;	Length 3646;
Best Local Similarity	100.0%;	Pred. No. 1.7e-197;			
Matches 786;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

QY	I	ATGGAGAGTCTGCCACGACAAAGAAATCTCTGACGACCCAGGTTGAATTCCTGTG	60
Db	3202	ATGGAGAGTCTGCCACGACAAAGAAATCTCTGACGACCCAGGTTGAATTCCTGTG	3143
QY	61	CAAGGCCAGGAGCCTGTGACCATGGAAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG	120
Db	3142	CAAGGCCAGGAGCCTGTGACCATGGAAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTG	3083
QY	121	GGCAGTTTCCCGCAGGTGGCCCGGCCGAGTGTGCTGAGACTCTCGGGGAGCCATTGACC	180
Db	3082	GGCAGTTTCCCGCAGGTGGCCCGGCCGAGTGTGCTGAGACTCTCGGGGAGCCATTGACC	3023
QY	181	ATCGTCTGTGAGATGGAGACTGGTGGACGGTGTGCTCAAGTCTCAGCAGAGAGTAT	240
Db	3022	ATCGTCTGTGAGATGGAGACTGGTGGACGGTGTGCTCAAGTCTCAGCAGAGAGTAT	2963
QY	241	AACATCCCCACGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC	300
Db	2962	AACATCCCCACGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGCGCTGAGC	2903
QY	301	AGGAGGAAAGCAGAGGAACTGTCTGTGTATCTGTGGAAACCTGGAGGGGGCCTTCTCATC	360
Db	2902	AGGAGGAAAGCAGAGGAACTGTCTGTGTATCTGTGGAAACCTGGAGGGGGCCTTCTCATC	2843
QY	361	CGGGAGAGCCAGACACGAGGAGGCTTACTCTGTGTAGTCCGCTCAGCCGCGCTTCA	420
Db	2842	CGGGAGAGCCAGACACGAGGAGGCTTACTCTGTGTAGTCCGCTCAGCCGCGCTTCA	2783
QY	421	TCCTGGAGCCGGATCAGACCTACAGAGATCACTAGCCTTTCACATATGGCTGGCTTTACATC	480
Db	2782	TCCTGGAGCCGGATCAGACCTACAGAGATCACTAGCCTTTCACATATGGCTGGCTTTACATC	2723
QY	481	TCACGGGCGCTCAACCTTCCCTCACTCCAGGCGCTGGTGGAGCCATTTACTCTGAAGCTGGCG	540
Db	2722	TCACGGGCGCTCAACCTTCCCTCACTCCAGGCGCTGGTGGAGCCATTTACTCTGAAGCTGGCG	2663
QY	541	GATGACATCTGTGCTTACTGAAAGAGCCCTGTGTCTGTGAGAGGGCTGGCCCGCTCCTT	600
Db	2662	GATGACATCTGTGCTTACTGAAAGAGCCCTGTGTCTGTGAGAGGGCTGGCCCGCTCCTT	2603
QY	601	GGCAAGGATATACCCCTTACCTGTGATCTGTGAGAGACACACATCAACTGSAAGAGCTG	660
Db	2602	GGCAAGGATATACCCCTTACCTGTGATCTGTGAGAGACACACATCAACTGSAAGAGCTG	2543
QY	661	GACAGCTCCCTCTCTTTTCTGAAAGCTGCCACAGGGAGAGAGTCTTTCTCAGTGAAGGT	720
Db	2542	GACAGCTCCCTCTCTTTTCTGAAAGCTGCCACAGGGAGAGAGTCTTTCTCAGTGAAGGT	2483
QY	721	CTCCGGAGTCCCTCAGCTTCTTACATCAAGCTGAATGACAGAGCTGTCTCTTTGATGATAT	780

PR 20-FEB-2001; 2001US-026961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (COR-) COR THERAPEUTICS INC.
XX Burgess CE, Conley PB, Groesse WM, Hart M, Kekuda R, Shinkets RA;
PI Sylek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX WPI; 2002-260937/32.
DR P-PSDB; AAU91308.
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Claim 1; Page 98; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albritght
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of various
CC human disorders, as well as in diagnostic applications. The present
CC sequence encodes a NOVX protein
XX
SQ Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;
Query Match 99.8%; Score 784.4; DB 6; Length 1183;
Best Local Similarity 99.9%; Pred. No. 3,2e-197;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 60
DB 398 ATGGGAAGTCTGCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTC 457
QY 61 CAAAGCCAGAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 458 CAAAGCCAGAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
QY 121 GGCAGATTTCGGGAGAGTGAGCCCGCGAGCTGTGCTGAGACTTGCGGAGCAATTAGC 180
DB 518 GGCAGATTTCGGGAGAGTGAGCCCGCGAGCTGTGCTGAGACTTGCGGAGCAATTAGC 577
QY 181 ATGCTCTCTGAGATGAGATGAGTGTGAGCGTGTCTGAAGTCTCAGGAGAGAGATAT 240
DB 578 ATGCTCTCTGAGATGAGATGAGTGTGAGCGTGTCTGAAGTCTCAGGAGAGAGATAT 637
QY 241 AACATCCCGACGCTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGAGCTGAGC 300
DB 638 AACATCCCGACGCTCCAGCTGGCCAAAGTCTCCCATGGGTGGCTGTATGAGGAGCTGAGC 697
QY 301 AGGAGGAAGCAGAGAACTGCTGTGTACCTGAGGAACCTTGAGGGGCTTCTCATC 360
DB 698 AGGAGGAAGCAGAGAACTGCTGTGTACCTGAGGAACCTTGAGGGGCTTCTCATC 757
QY 361 CGGAGAGCCAGACCAAGAGAGGCTTACTCTCTGTAGTCCGCTAGCCGCCCTTCA 420

DB 758 CGGAGAGCCAGACCAAGAGAGGCTTACTCTCTGTCACTCCGCTCAGCCGCCCTGCA 817
QY 421 TCCCTGGAGCCGATCAGACACTACAGAGATCCAGCTTGTGACATGAGCTGATACATC 480
DB 818 TCCCTGGAGCCGATCAGACACTACAGAGATCCAGCTTGTGACATGAGCTGATACATC 877
QY 481 TCACCGGCTCCACCTTCCCTCACTCAGAGCCCTGTGAGCACTTACTGAGCTGAGC 540
DB 878 TCACCGGCTCCACCTTCCCTCACTCAGAGCCCTGTGAGCACTTACTGAGCTGAGC 937
QY 541 GATGACATCTGCTGCTCCTCACTCAAGAGACCTGTGCTCTGAGAGAGGCTGCGCTCCT 600
DB 938 GATGACATCTGCTGCTCCTCACTCAAGAGACCTGTGCTCTGAGAGAGGCTGCGCTCCT 997
QY 601 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGAGAAAGACTG 660
DB 998 GGCAGAGATATACCCCTACCTGTGACTGTGACAGAGACACCACTCAACTGAGAAAGACTG 1057
QY 661 GACAGCTCCCTCTGCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTTCTGAGTGAAGGT 720
DB 1058 GACAGCTCCCTCTGCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTTCTGAGTGAAGGT 1117
QY 721 CTCGGGAGTCCCTCAGCTCTACATGAGCTGAAAGAGAGGCTGCTCTTGGAGAT 780
DB 1118 CTCGGGAGTCCCTCAGCTCTACATGAGCTGAAAGAGAGGCTGCTCTTGGAGAT 1177
QY 781 GCCTAG 786
DB 1178 GCCTAG 1183
RESULT 7
AAC77202
ID AAC77202 standard; cDNA; 837 BP.
XX
AC AAC77202;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF2757 polynucleotide sequence SEQ ID NO:5513.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiposrotatic; antiparkinsonian; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antihypertensive; immunosuppressant; cardiac;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
KW antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antineoplastic disease; coagulation;
KW thrombosis; contraceptive; ss.
XX
OS Homo sapiens.
XX
XX
PN MO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000OWO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Shinkets RA, Leach M;
XX

ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. hematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC

XX Sequence 1413 BP; 332 A; 407 C; 388 G; 286 T; 0 U; 0 Other;

Query Match 85.4%; Score 672.8; DB 6; Length 1413;
 Best Local Similarity 93.4%; Pred. No. 1,1e-167;
 Matches 734; Conservative 0; Mismatches 2; Indels 50; Gaps 1;

```

QY 1 ATGGGAAGTCTGCCCGACGAGAAATCTCTGCCAAGCCCAAGCTTGAATTCCTCTGTC 60
DB |||||
QY 54 ATGGGAGTCTGCCCGACGAGAAATCTCTGCCAAGCCCAAGCTTGAATTCCTCTGTC 113
DB |||||
QY 61 CAAGGCCAGGACCTGTGACCATGAAAGCAAGAGCAAGGCCACAGCCGTCGCTG 120
DB |||||
QY 114 CAAGGCCAGGACCTGTGACCATGAAAGCAAGAGCAAGGCCACAGCCGTCGCTG 173
DB |||||
QY 121 GGCAGTTTCCCGGAGGTGGCCCGGACGCTGTGAGACTCGGGGAGCCATTGACC 180
DB |||||
QY 174 GGCAGTTTCCCGGAGGTGGCCCGGACGCTGTGAGACTCGGGGAGCCATTGACC 223
DB |||||
QY 181 ATGCTCTTGAAGATGAGACTGGTGGACGCTGTGAGAGCTGAGGACGAGAGATAT 240
DB |||||
QY 234 ATGCTCTTGAAGATGAGACTGGTGGACGCTGTGAGAGCTGAGGACGAGAGATAT 293
DB |||||
QY 241 AACATCCCGACGCTGCAGTGGCCAAAGTCTCCCATGAGTGGCTGTATGAGGGCTGAGC 300
DB |||||
QY 294 AACATCCCGACGCTGCAGTGGCCAAAGTCTCCCATGAGTGGCTGTATGAGGGCTGAGC 353
DB |||||
QY 301 AGGGAAGAAAGCAAGGAACTGCTGTGTACTTCTGAGAACTTGGAGGGGCTTCTCATC 360
DB |||||
QY 354 AGGGAAGAAAGCAAGGAACTGCTGTGTACTTCTGAGAACTTGGAGGGGCTTCTCATC 413
DB |||||
QY 361 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGCA 420
DB |||||
QY 414 CGGAGAGCCAGACACAGAGAGGCTCTTACTCTGTGAGTCCGCTCAGCGCCCTGCA 473
DB |||||
QY 421 TCCTGGAGCCGATCAGACACTACAGAGTCCACTGCTTGAACAATGGCTGTATCATC 480
DB |||||
QY 474 TCCTGGAGCCGATCAGACACTACAGAGTCCACTGCTTGAACAATGGCTGTATCATC 533
DB |||||
QY 481 TCACCGGCTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
DB |||||
QY 534 TCACCGGCTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 584
DB |||||
QY 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGCTGAGAGGCTGAGGCTGCTCTCT 600
DB |||||
QY 585 -----GAGGGCTGGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 603
DB |||||
QY 601 GCGAAGATATACCTTACCTGTGACTGTGACAGAGACACCACTCACTGAAAGAGCTG 660
DB |||||
QY 604 GCGAAGATATACCTTACCTGTGACTGTGACAGAGACACCACTCACTGAAAGAGCTG 663
DB |||||
QY 661 GACAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
DB |||||
QY 664 GACAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 723
DB |||||
QY 721 CTCGGGAGTCCCTGAGCTTCTACATCAGCTGAAATGACAGAGCTGTCTCTTTGGATAT 780
DB |||||
QY 724 CTCGGGAGTCCCTGAGCTTCTACATCAGCTGAAATGACAGAGCTGTCTCTTTGATAT 783
DB |||||
QY 781 GCTTAG 786
DB |||||

```

Db 784 GCCTAG 789

RESULT 9
 AAL44090
 ID AAL44090 standard; cDNA; 737 BP.

AC AAL44090;
 XX
 DT 03-OCT-2002 (first entry)

DE Mouse MARS short isoform protein coding sequence.

XX Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;
 KM MARS; tumour suppressor gene; Scr-like adaptor protein; SHAP;
 KM myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
 XX immunosuppression; myeloproliferative disorder; breast cancer.

OS Mus sp.

XX Key Location/Qualifiers
 FH 1..633
 FT CDS /tag= a
 FT /product= "Mouse MARS short isoform protein"

XX MO200242452-A2.

XX PD 30-MAY-2002.

XX PF 26-NOV-2001; 2001MO-CA001662.

XX PR 27-NOV-2000; 2000CA-02324663.

XX PA (HOSP-) HOSPITAL FOR SICK CHILDREN.

XX PI Meglade JC, Loreto MP;

XX XX WPI; 2002-566564/60.

XX DR P-PSDB; AAO15458.

XX PT New isolated modulator of antigen receptor signalling protein or its
 PT fragment, useful for treating malignant disorders such as myeloid
 PT malignancies, autoimmune disorders and myeloproliferative disorders.

XX PS Claim 9; Page 77; 110bp; English.

CC The invention comprises the amino acid and coding sequences of modulator
 CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
 CC putative tumour suppressor gene and exhibits structural and sequence
 CC similarity to the Scr-like adaptor protein (SHAP). The MARS DNA and
 CC protein sequences of the invention are useful for the treatment of
 CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
 CC disorders, immunosuppression, myeloproliferative disorders and
 CC malignancies related to the de-regulation of tyrosine kinases (e.g.
 CC breast cancer). The present cDNA sequence encodes a mouse MARS protein
 CC

XX Sequence 737 BP; 152 A; 219 C; 218 G; 148 T; 0 U; 0 Other;

Query Match 84.4%; Score 663.4; DB 6; Length 737;
 Best Local Similarity 93.4%; Pred. No. 2.7e-165;
 Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

```

QY 1 ATGGGAAGTCTGCCCGACGAGAAATCTCTGCCAAGCCCAAGCTTGAATTCCTCTGTC 60
DB |||||
QY 61 CAAGGCCAGGACCTGTGACCATGAAAGCAAGAGCAAGGCCACAGCCGTCGCTG 120
DB |||||
QY 61 CAAGGCCAGGACCTGTGACCATGAAAGCAAGAGCAAGGCCACAGCCGTCGCTG 120
DB |||||
QY 61 CAAGGCCAGGACCTGTGACCATGAAAGCAAGAGCAAGGCCACAGCCGTCGCTG 120
DB |||||
QY 121 GGCAGTTTCCCGGAGGTGGCCCGGACGCTGTGAGACTCGGGGAGCCATTGACC 180
DB |||||
QY 121 GGCAGTTTCCCGGAGGTGGCCCGGACGCTGTGAGACTCGGGGAGCCATTGACC 180
DB |||||

```

```

QY 181 ATGCTCTGAGGATGAGACGTGTGGAAGCTCTGTTAAAGTCTCAGGCGAGAGTAT 240
Db 181 ATGCTCTGAGGATGAGACGTGTGGAAGCTCTGTTAAAGTCTCAGGCGAGAGTAT 240
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGAGCTGTATAGAGGCTGAGC 300
Db 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGAGCTGTATAGAGGCTGAGC 300
QY 301 AGGAGAAAGCAGAGGAACCTGTTGTTAACTGGGAACTCTGAGAGGCTCTCCCTCATC 360
Db 301 AGGAGAAAGCAGAGGAACCTGTTGTTAACTGGGAACTCTGAGAGGCTCTCCCTCATC 360
QY 361 CGGAGAGCCAGACCAAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
Db 361 CGGAGAGCCAGACCAAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
QY 421 TCCTGGAGCCGAGTCAAGACTACAGAGATCCAGCTCTGACATGAGGCTGATACATC 480
Db 421 TCCTGGAGCCGAGTCAAGACTACAGAGATCCAGCTCTGACATGAGGCTGATACATC 480
QY 481 TCACCGCGCTCACCTTCCCTCACTCCAGGCTGTGGACCATTAATCT 531
Db 481 TCACCGCGCTCACCTTCCCTCACTCCAGGCTGTGGACCATTAATCT 531
QY 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTGACAGAGGCTGGCCGCTCCCT 600
Db 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTGACAGAGGCTGGCCGCTCCCT 600
QY 601 GGCAAGATATACCCCTACTGTGACTGTGACAGAGACCACTCACTGAGAAAGAGTGT 660
Db 601 GGCAAGATATACCCCTACTGTGACTGTGACAGAGACCACTCACTGAGAAAGAGTGT 660
QY 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGGAGAGAGTCTTCTGATGAGAGT 720
Db 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGGAGAGAGTCTTCTGATGAGAGT 720
QY 721 CTCGGGAGTCCCTGAGCTTCAATCACTGAGCTGATGA-CAAGGTGTCTCTTGGATGA 779
Db 721 CTCGGGAGTCCCTGAGCTTCAATCACTGAGCTGATGA-CAAGGTGTCTCTTGGATGA 779
QY 780 TGCCTAG 786
Db 780 TGCCTAG 786
QY 731 TGCCTAG 737
Db 731 TGCCTAG 737

```

RESULT 10

```

AAS74750
ID AAS74750 standard; cDNA; 2049 BP.
XX
AC AAS74750;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10554.
XX
KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;

```

```

XX WPI; 2001-639362/73.
DR P-PSDB; ABG10563.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 10554; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp://wipo.int/pub/published_pct_sequences

SQ
Sequence 2049 BP; 479 A; 573 C; 551 G; 443 T; 0 U; 3 Other;

Query Match 77.6%; Score 609.8; DB 5; Length 2049;
Best Local Similarity 92.2%; Pred. No. 5.6e-15;
Matches 718; Conservative 0; Mismatches 7; Indels 54; Gaps 5;

QY 1 ATGGGAAGTCTGCCAGAGAAAGAAATCTTGCCAGCCCAAGTTGATTCCTCTGTC 60
Db 965 ATGGGAAGTCTGCCAGAGAAAGAAATCTTGCCAGCCCAAGTTGATTCCTCTGTC 1024
QY 61 CAAGGCCAGAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 1025 CAAGGCCAGAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1084
QY 121 GGCAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGACCAATTGACC 180
Db 1085 GGCAGTTTCCCGGAGGTGGCCCGGAGCTGTGCTGAGACTCGGGAGACCAATTGACC 1144
QY 181 ATGCTCTGAGGATGAGAGCTGTGAGACGTTGCTGTAAGTCTCAGGAGAGAGTAT 240
Db 1145 ATGCTCTGAGGATGAGAGCTGTGAGACGTTGCTGTAAGTCTCAGGAGAGAGTAT 1204
QY 241 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGAGCTGTATAGAGGCTGAGC 300
Db 1205 AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGGAGCTGTATAGAGGCTGAGC 1264
QY 301 AGGAGAAAGCAGAGGAACCTGTTGTTAACTGGGAACTCTGAGAGGCTCTCCCTCATC 360
Db 1265 AGGAGAAAGCAGAGGAACCTGTTGTTAACTGGGAACTCTGAGAGGCTCTCCCTCATC 1324
QY 361 CGGAGAGCCAGACCAAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
Db 1325 CGGAGAGCCAGACCAAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCTGCA 1384
QY 421 TCCTGGAGCCGAGTCAAGACTACAGAGATCCAGCTCTGACATGAGGCTGATACATC 480
Db 1385 TCCTGGAGCCGAGTCAAGACTACAGAGATCCAGCTCTGACATGAGGCTGATACATC 1444
QY 481 TCACCGGCTCACCTTCCCTCACTCCAGGCTGTGGACCATTAATCTGAGCTGAGC 540

```

Dd	1445	TCAACGGGCGCTCACCTTCCCTACTCAGGCCCTGTGGAGCATATTA	CTCT-----	1495S
Qy	541	GATGACATCTGCTGCTTACTTCAGAGGCCCTGTGTCTCTGCAGAGGGGCTGGCCGCTCCCT		600
Dd	1496	-----	GAAGGGCTGGGCCCTCCCT	1514
Qy	601	GGCAAGATATACCCTTACCTTGATGCTGTGAGAGGACACACCACTCAACTGGAAAGAGCTG		660
Dd	1515	GGCAAGATATACCCTTACCTTGATGCTGTGAGAGGACACCACTCAACTGGAAAGAGCTG		1574
Qy	661	GACAGCTCCCTCTGTTTCTTGAAGCTGCCACA-GGGAGAGAGTCTCTTCTCAG-TGAGG		718
Dd	1575	GACAGCTCCCTCTGTTTCTTGAAGCTGCCACAAGGGAGAGAGTCTCTTCTCAGAGAGG		1634
Qy	719	GTCCTCCGGAGT-CCCTCAGCTTTACTATCAG-CTGAATGACGAGGCTGTCTCTTTGG		775
Dd	1635	GGCTCCGGAGTCCCTCAGCTTTACTATCAGCCTGAATGACGAGGCTGTCTCTTTTG		1693

RESULT 11			
AAL44087	ID	AAL44087 standard; cDNA; 1348 BP.	
XX	AC	AAL44087;	
DT	03-OCT-2002	(first entry)	
XX	Mouse modulator of antigen receptor signalling protein coding sequence.		
KW	Mouse; gene; ss; gene therapy; modulator of antigen receptor signalling;		
KM	MARS; tumour suppressor gene; Scr-like adaptor protein; SLAP;		
KX	myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;		
XX	immunosuppression; myeloproliferative disorder; breast cancer.		
OS	Mus sp.		
FH	Key	Location/Qualifiers	
FT	CDS	282..1061	
FT		/tag= a	
FT		/product= "Mouse MARS protein"	
XX	PN	MO200242452-A2.	
PD	30-MAY-2002.		
XX	26-NOV-2001;	2001WO-CA001662.	
Pf	27-NOV-2000;	2000CA-02324663.	
PR	(HOSE-) HOSPITAL FOR SICK CHILDREN.		
XX	Mcglade JC, Loreto MP;		
PI	DR	WPI: 2002-566564/60.	
PT	DR	P-PsDB: AA015456.	
XX	New isolated modulator of antigen receptor signalling protein or its fragment, useful for treating malignant disorders such as myeloid malignancies, autoimmune disorders and myeloproliferative disorders.		
XX	Claim 10; Fig 1A; 11opp; English.		
XX	The invention comprises the amino acid and coding sequences of modulator of antigen receptor signalling (MARS) proteins. The MARS protein is a putative tumour suppressor gene and exhibits structural and sequence similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and protein sequences of the invention are useful for the treatment of myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune disorders, immunosuppression, myeloproliferative disorders and malignancies related to the de-regulation of tyrosine kinases (e.g. breast cancer). The present cDNA sequence encodes a mouse MARS protein		
SQ	Sequence 1348 BP; 324 A; 385 C; 362 G; 277 T; 0 U; 0 Other;		

Query Match	68.4%;	Score 537.8;	DB 6;	Length 1348;
Best Local Similarity	81.9%;	Pred. No. 5.4e-132;		
Matches 646;	Conservative 0;	Mismatches 137;	Indels 6;	Gaps 2

QY	1	ATGGGAAGTGTGGCCGAGCAAGAAATAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTGC	60
Db	282	ATGGGAAGTTTGTTCAGCAGAGGGAATAAC---CTCAGCCCCAGCCCAAGCTCTCTGTGT	338
QY	61	CAAGGCCAGGACCTTGACCATGGAAGCAGAGAAAGCAAGGCCACAGCCGTGGCCCTTG	120
Db	339	CCAGACCAGGAACCCGCTGTCCATGCAACCGAATAAGACACAGGTGTCAAGCTGTGGCCCTG	398
QY	121	GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTGCTGAGACTGGGGAGACCATTGACC	180
Db	399	GGCAGTTTCCACAGGTGAACAGGCCACGACTACTCTTGAGACTCGGGAGCCGCTTGACC	458
QY	181	ATCGTCTCTAGGATGAGACTGGTGGACCGGTGCTGTCTGAACTCTCAGGCAGAGATAT	240
Db	459	ATCATCTCTAGATGAGATGGTAAGTTGGTGGACAGTCCAGTCGAAAGTCTCAGGCAGAGATAC	518
QY	241	AACATCCCCAGCGTCCACGTGGCCAAAGTCTCCCATGGTGGCTGTATGAGGGCTGAGC	300
Db	519	CACATGCCCAAGTGTGTGTGTGGCTAAAGTGGCCACGGGTGGCTGTATCGAGGGCTTGAGC	578
QY	301	AGGAGAAAGCAGAGAGAACTGCTGTGTGTAACCTGGGAACCTGAGAGGGGCTTCTCATC	360
Db	579	CGGAGAAAGCCGAGGAACCTACTCTTTACTGTGGAAACCCCGAGGGGCTTCTCATC	638
QY	361	CGGAGAGCCAGACCCAGAGAGGCTCTTACTCTGTCACTGTCCGCTCAGCCGCCCTGCA	420
Db	639	CGGAGAGCCAGACCCAGAGAGGCTGTGATATCCCTGTCCGCTCAGCTCAGCCGCCCTGCA	698
QY	421	TCTTGGGACCGGATACAGCACTAAGAGATCCACTGCTGCTTGAACAATGGCTGGCTGTATC	480
Db	699	TCTTGGGACCGGATACAGCACTAAGAGATAGCGTCTTGACAAATGGCTGGCTGTATC	758
QY	481	TCACCGGCGCTCACTTCCCTCACTCCAGGCGCTGGTGGACCATTACTGTGACTGGCG	540
Db	759	TCACCTGGCTCACTTCCCTCACTCCAGCGCTTGGTGGAGATTACTGTGACTGACA	818
QY	541	GATGACATCTGCTGCTACTCAAGAGCGCTGTGTCTGTGAGAGGCTGCGCCGCTCCTT	600
Db	819	GATGGACATCTGCTGTGCCCTCAGGAGCGCGTGTGCTGTGACGAAGCTTGGGCCACTACTT	878
QY	601	GGCAAGGATTAACCCCTAACCCTGATACGTGACAGAGGACACACACTCAACTGSAAGAAGCTG	660
Db	879	GGCAAGGATTAACCTCCACCTGTGACTGTGGCAATCATATACTAAATGSAAAAAGCTG	938
QY	661	GACAGCTCCCTCTGTGTTTCTGAAG--CTGACCAGGGGAGAGACTCTTCTCAGTAG	717
Db	939	GACCGACACTCTCTGTCTTGGAAGCACTGCGAGTGGGAGGACATCTCTCTCAGTAG	998
QY	718	GCTCTCCGGAGTCCCTCAGCTTCTACATCAGCCTGAAATGACGAGGCTGTCTCTTGGAT	777
Db	999	GAGCTCCAGAGTCCCTCAGTCTCAATCAGCCTGCTGAGGACCCCTTGATGATGCT	1058
QY	778	GATGCTAG 786	
Db	1059	TAGCCCTG 1067	

RESULT 12
ABQ98670
ID ABQ98670 standard; DNA; 763 BP.
XX
XX
AC ABQ98670;
XX
DT 04-NOV-2002 (first entry)
XX
XX
DE Human ORF477 coding sequence.
XX
XX
KW Cytostatic; Carcinant; Anti-allergic; Immunosuppressive; Vulnerary;

KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
 KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
 KW cancer; cardiovascular disease; allergy; autoimmune disease;
 KW wound healing; blood coagulation disorder; inflammatory disorder; ds.

OS Homo sapiens.

PN US2002082206-A1.

PD 27-JUN-2002.

PF 30-MAY-2001; 2001US-00867550.

PR 30-MAY-2000; 2000US-0208427P.

PA (LEACH/) LEACH M D.

PA (MEHRABAN F.) MEHRABAN F.

PA (CONLEY P B.) CONLEY P B.

PA (TOPPER J N.) TOPPER J N.

PA (LAW D.) LAW D.

PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;

DR WPI; 2002-626554/67.

DR P-PSDB; ABP64107.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Claim 2; SEQ ID NO 953; 78pp; English.

The present invention relates to novel human ORFX polypeptides and their coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences were discovered in human atherogenic cells, in particular in platelets and human umbilical vein endothelial cells (HUVEC) and are expressed in many other tissues as well. Atherogenic cells are cells which have the potential to develop atherosclerotic plaques. The ORFX polypeptides and nucleic acids are useful for treating or preventing a pathological condition associated with an ORFX-associated disease, e.g. cancer, cardiovascular disease, allergy, autoimmune disease, wound healing, blood coagulation disorders or inflammatory disorders. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at seqdata.uspto.gov/sequence.html?docID=20020082206

Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 60.8%; Score 478; DB 6; Length 763;

Best Local Similarity 100.0%; Pred. No. 3e-116;

Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGAGAAATCTTGTGCAAGCCCAAGCTTGAATCTCTGTC 60

Db 286 ATGGGAAGTCTGCCAGAGAAATCTTGTGCAAGCCCAAGCTTGAATCTCTGTC 345

QY 61 CAAGGCCAGGAGCTGTGACCATGGAAGCAAGGCAAGGCGGAGCCCTG 120

Db 346 CAAGGCCAGGAGCTGTGACCATGGAAGCAAGGCAAGGCGGAGCCCTG 405

QY 121 GGCAGTTCCCGGAGGTGGCCGGCGAGTGTGCTGAGACTCGGGAGCCATTGACC 180

Db 406 GGCAGTTCCCGGAGGTGGCCGGCGAGTGTGCTGAGACTCGGGAGCCATTGACC 465

QY 181 ATGCTCTAGAGATGAGACTGTGACCGTGTCTGAAGCTCAGGCAAGAGTAT 240

Db 466 ATGCTCTAGAGATGAGACTGTGACCGTGTCTGAAGCTCAGGCAAGAGTAT 525

QY 241 AATATCCCGGAGGTCCAGCTGTCCTCCATGGTGGCTGTATGAGGGGCTTGAAC 300

Db 526 AATATCCCGGAGGTCCAGCTGTCCTCCATGGTGGCTGTATGAGGGGCTTGAAC 585

QY 301 AGGAGAAAGACAGAGAACTGCTGTTACTCTGGAAACCTGAGAGGGGCTTCCATAC 360

Db 586 AGGAGAAAGACAGAGAACTGCTGTTACTCTGGAAACCTGAGAGGGGCTTCCATAC 645

QY 361 CGGAGAGCCAGACCAAGAGAGCTTACTCTGTCAATCCGCTCAAGCCGCTTGA 420

Db 646 CGGAGAGCCAGACCAAGAGAGCTTACTCTGTCAATCCGCTCAAGCCGCTTGA 705

QY 421 TCCTGGACCCGATGACACTACAGATCCACCTGTAACATGGCTGCTGATACA 478

Db 706 TCCTGGACCCGATGACACTACAGATCCACCTGTAACATGGCTGCTGATACA 763

RESULT 13

AA574748 ID AA574748 standard; cDNA; 603 BP.

AA574748;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #10552.

Human; chromosome mapping; gene mapping; gene therapy; forensic;

Food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG10561.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits and to assess biodiversity.

Claim 1; SEQ ID NO 10552; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain

reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used

in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal

activity of (III) or to treat disease states involving (II). (II) is

useful for generating antibodies against it, detecting or quantitating a

polypeptide in tissue, as molecular weight markers and as a food

supplement. (II) and its binding partners are useful in medical imaging

of sites expressing (II). (I) and (II) are useful for treating disorders

involving aberrant protein expression or biological activity. The

diagnostics, forensics, gene mapping, identification of mutations

responsible for genetic disorders or other traits to assess biodiversity

and to produce other types of data and products dependent on DNA and

amino acid sequences. AA564197-AA594564 represent novel human diagnostic

coding sequences of the invention. Note: The sequence data for this

patent did not appear in the printed specification, but was obtained in

electronic format directly from WIPO at

ftp.wipo.int/pub/published_pat_sequences

This Page Blank (uspto)

QY 241 AACATCCCAAGGCTTCACGTTGACCAAGTCTCCATGGGAGCTGTATGAGGGCTGAGC 300
 DB 241 AACATCCCAAGGCTTCACGTTGACCAAGTCTCCATGGGAGCTGTATGAGGGCTGAGC 300
 QY 301 AGGAGAAAGCAGAGGAACTGTGTTTGAACCTGGAGACCTTGAGAGGCTTCTCTATC 360
 DB 301 AGGAGAAAGCAGAGGAACTGTGTTTGAACCTGGAGACCTTGAGAGGCTTCTCTATC 360
 QY 361 CGGAGAGCCAGACCGAGAGAGGCTCTTACTCTCTGCAATCCGCTCAAGCCGCTGCA 420
 DB 361 CGGAGAGCCAGACCGAGAGAGGCTCTTACTCTCTGCAATCCGCTCAAGCCGCTGCA 420
 QY 421 TCCTGGAGACCGATTCAGACTACAGGATCCATGCTGACATGGCTGAGTGGCTGATC 480
 DB 421 TCCTGGAGACCGATTCAGACTACAGGATCCATGCTGACATGGCTGAGTGGCTGATC 480
 QY 481 TCACCGGCTCTCACTTCCCTCACTTCAGAGCCCTGGTGGACCATTACTGTAGCTGGC 540
 DB 481 TCACCGGCTCTCACTTCCCTCACTTCAGAGCCCTGGTGGACCATTACTGTAGCTGGC 540
 QY 541 GATGACATCTGCTGCTCTACTCAAGAGCCCTGTGTCTTGAGAGGCTGGCCGCTCCCT 600
 DB 541 GATGACATCTGCTGCTCTACTCAAGAGCCCTGTGTCTTGAGAGGCTGGCCGCTCCCT 600
 QY 601 GGCAGAGATATACCCCTACCTGATGATGTCGACAGAGACACCACTCACTGAGAAAGAGCTG 660
 DB 601 GGCAGAGATATACCCCTACCTGATGATGTCGACAGAGACACCACTCACTGAGAAAGAGCTG 660
 QY 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTCTTCTGAGAGGT 720
 DB 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTCTTCTGAGAGGT 720
 QY 721 CTCGGAGAGTCCCTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 721 CTCGGAGAGTCCCTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 781 GCCTAG 786
 DB 781 GCCTAG 786

RESULT 2
 AX572845 786 bp DNA linear PAT 29-NOV-2002
 LOCUS Sequence 1 from Parent W002055707.
 AX572845
 ACCESSION AX572845
 VERSION AX572845.1 GI:26004935
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C.,
 Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M.,
 Liao, X.C. and Speng, N.
 Cloning of an inhibitor of antigen-receptor signaling by a
 retroviral-based functional screen
 Patent: WO 02055707-A 1 18-JUL-2002;
 Rigel Pharmaceuticals, Inc. (US)
 Location/Qualifiers
 1..786
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 1..786
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD58542.1"
 /db_xref="GI:26004935"
 /translation="MGSLPSRRKSLPSPLSSSSVQGGPVTMEARSKATVALGSPF
 AGPAELSLRGEPLTIVSEDDWMTVLSEVSGREYNIPVHAKVSHGLYEGLSRE

CDS
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD58542.1"
 /db_xref="GI:26004935"
 /translation="MGSLPSRRKSLPSPLSSSSVQGGPVTMEARSKATVALGSPF
 AGPAELSLRGEPLTIVSEDDWMTVLSEVSGREYNIPVHAKVSHGLYEGLSRE

ORIGIN

KAEELLLPANGAFILRESOTRRGSIYSLSVLSRPSASWDIRIRHICLNDGMLYI
 SPRLTFPSLQALVDHYSLELADIICLILKPCVLAGAPLPGKIDPLPVTQPTPLNWK
 HEDSLTFSEATGSESLISEGLRESISFYISINDENVSLIDA"

Query Match 100.0%; Score 786; DB 6; Length 786;
 Best Local Similarity 100.0%; Pred.No. 8,5e-195;
 Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGCTGCGCCAGCAAGAAATCTCTCCAAAGCCCAAGCTGAGTTCTCTGTC 60
 DB 1 ATGGGAAGCTGCGCCAGCAAGAAATCTCTCCAAAGCCCAAGCTGAGTTCTCTGTC 60
 QY 61 CAAGCCAGAGGACCTTGTGAACATGAGAACAGAGAAAGCAAGCCACAGCTGGGCTTG 120
 DB 61 CAAGCCAGAGGACCTTGTGAACATGAGAACAGAGAAAGCAAGCCACAGCTGGGCTTG 120
 QY 121 GGCAGTTTCCCGCAGAGTGCCCGGCCAGACTCTCGTGAATCTGGGGAGCCATTGACC 180
 DB 121 GGCAGTTTCCCGCAGAGTGCCCGGCCAGACTCTCGTGAATCTGGGGAGCCATTGACC 180
 QY 181 ATCGTCTGAGAGATGAGACTGATGACAGGTCTGTCTGAAGTCTGAGGACAGAGATAT 240
 DB 181 ATCGTCTGAGAGATGAGACTGATGACAGGTCTGTCTGAAGTCTGAGGACAGAGATAT 240
 QY 241 AACATCCCAAGGCTTCACGTTGACCAAGTCTCCATGGGAGCTGTATGAGGGCTGAGC 300
 DB 241 AACATCCCAAGGCTTCACGTTGACCAAGTCTCCATGGGAGCTGTATGAGGGCTGAGC 300
 QY 301 AGGAGAAAGCAGAGGAACTGTGTTTGAACCTGGAGACCTTGAGAGGCTTCTCTATC 360
 DB 301 AGGAGAAAGCAGAGGAACTGTGTTTGAACCTGGAGACCTTGAGAGGCTTCTCTATC 360
 QY 361 CGGAGAGCCAGACCGAGAGAGGCTCTTACTCTGTGACGTGGCCCTCAAGCCGCTGCA 420
 DB 361 CGGAGAGCCAGACCGAGAGAGGCTCTTACTCTGTGACGTGGCCCTCAAGCCGCTGCA 420
 QY 421 TCCTGGAGACCGATTCAGACTACAGGATCCATGCTGACATGGCTGAGTGGCTGATC 480
 DB 421 TCCTGGAGACCGATTCAGACTACAGGATCCATGCTGACATGGCTGAGTGGCTGATC 480
 QY 481 TCACCGGCTCTCACTTCCCTCACTTCAGAGCCCTGGTGGACCATTACTGTAGCTGGC 540
 DB 481 TCACCGGCTCTCACTTCCCTCACTTCAGAGCCCTGGTGGACCATTACTGTAGCTGGC 540
 QY 541 GATGACATCTGCTGCTCTACTCAAGAGCCCTGTGTCTTGAGAGGCTGGCCGCTCCCT 600
 DB 541 GATGACATCTGCTGCTCTACTCAAGAGCCCTGTGTCTTGAGAGGCTGGCCGCTCCCT 600
 QY 601 GGCAGAGATATACCCCTACCTGATGATGTCGACAGAGACACCACTCACTGAGAAAGAGCTG 660
 DB 601 GGCAGAGATATACCCCTACCTGATGATGTCGACAGAGACACCACTCACTGAGAAAGAGCTG 660
 QY 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTCTTCTGAGAGGT 720
 DB 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCACAGGGAGAGAGTCTCTTCTGAGAGGT 720
 QY 721 CTCGGAGAGTCCCTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 721 CTCGGAGAGTCCCTCACTCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 781 GCCTAG 786
 DB 781 GCCTAG 786

RESULT 3
 AF290985 786 bp mRNA linear PRI 21-JAN-2003
 LOCUS Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
 AF290985
 ACCESSION AF290985
 VERSION AF290985.1 GI:17351920
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 786)
 Loreto, M.P. and McGlade, C.J.
 Cloning and characterization of human Src-like adaptor protein 2
 and a novel splice isoform, SLAP-2-v
 Oncogene 22 (2), 266-273 (2003)
 JOURNAL MEDLINE 22415750
 PUBMED 12527895
 REFERENCE 2 (bases 1 to 786)
 Loreto, M.P. and McGlade, C.J.
 Direct Submission
 Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
 Location/Qualifiers
 1..786
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /tissue_type="thymus"
 1..786
 /note="SLAP-2"
 /codon_start=1
 /product="Src-like adaptor protein-2"
 /protein_id="AAJ3197.1"
 /db_xref="GI:17351921"
 /translation="MGSLLPSSRRKSLPSPSSVQGGPVTMEERSKATVALGSP
 AGPAELSLRGEPLTIVSDGDMWTLSVSGREYNI PSYHAKVSHGWLVEGLSRE
 KAELLILPBNPGAFILIRSDQTRGYSLSVRLSPRSMIRIHYRIHCLDNGMLYI
 SPRLTPPSIQALVHYSHLADIDICLLKEPCVLRAPGLPKDIPLPVTQRPPLNWK
 ELDSLLFSEATGESLISRLRESLSFYISLNDVSLDDA"
 ORIGIN
 Query Match 100.0%; Score 786; DB 9; Length 786;
 Best Local Similarity 100.0%; Pred. No. 8.5e-195;
 Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGGAAGCTGCTCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTC
 DB 1 ATGGGAAGTGTCCAGAGAAATCTCTGCCAAGCCCAAGCTTGAAGTCTCTGTC 60
 QY 61 CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAAAGCAAGCCGCGTGCCTG 120
 DB 61 CAAGGCCAGGAGCTGTGACCATGGAAGCAGAGAAAGCAAGCCGCGTGCCTG 120
 QY 121 GGCAGTTTCCGCGAGGTGGCCGCGAGCTGTGCTGAGACTCCGGGAGCCATTGACC 180
 DB 121 GGCAGTTTCCGCGAGGTGGCCGCGAGCTGTGCTGAGACTCCGGGAGCCATTGACC 180
 QY 181 ATCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 240
 DB 181 ATCTCTCTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 240
 QY 241 AACATCCCGAGGCTTCCAGTGGCCAAAGTCTCCATGAGTGGCTGTATGAGGCGCTGAGC 300
 DB 241 AACATCCCGAGGCTTCCAGTGGCCAAAGTCTCCATGAGTGGCTGTATGAGGCGCTGAGC 300
 QY 301 AGGAGAAACAGAGAAAGTCTGTGTTACCTGGGAACCTGGAGGGGCTTCCTCATC 360
 DB 301 AGGAGAAACAGAGAAAGTCTGTGTTACCTGGGAACCTGGAGGGGCTTCCTCATC 360
 QY 361 CGGAGAGCCAGACAGAGAGAGCTTACTCTGTGAGTCCGCTCAGCCGCGCTGCA 420
 DB 361 CGGAGAGCCAGACAGAGAGAGCTTACTCTGTGAGTCCGCTCAGCCGCGCTGCA 420
 QY 421 TCCGAGGAGCCAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 480
 DB 421 TCCGAGGAGCCAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 480
 QY 481 TCACCGGCTCACTTCCCTCACTCAAGCCCTGTGAGCACTTACTGTGAGCTGGCG 540

DB 481 TCACCGGCTCACTTCCCTCACTCAAGCCCTGTGAGCACTTACTGTGAGCTGGCG 540
 QY 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTTCGAGAGGGCTGCGCTCCCT 600
 DB 541 GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGCTTCGAGAGGGCTGCGCTCCCT 600
 QY 601 GGCAGGATATACCCCTACTGTGACTGTGACAGAGAGACCACTCACTGAAAGAGCTG 660
 DB 601 GGCAGGATATACCCCTACTGTGACTGTGACAGAGAGACCACTCACTGAAAGAGCTG 660
 QY 661 GACAGCTTCCCTGCTGTTTCTGAAAGCTGCCACAGGAGAGAGTCTTCTCAGTAGAGT 720
 DB 661 GACAGCTTCCCTGCTGTTTCTGAAAGCTGCCACAGGAGAGAGTCTTCTCAGTAGAGT 720
 QY 721 CTCGGGAGTCCCTCAGCTTCAATGAGCCCTGAAATGACAGAGCTGCTCTTGATGAT 780
 DB 721 CTCGGGAGTCCCTCAGCTTCAATGAGCCCTGAAATGACAGAGCTGCTCTTGATGAT 780
 QY 781 GCCTAG 786
 DB 781 GCCTAG 786
 RESULT 4
 AF326353
 LOCUS 786 bp mRNA linear PRI 08-NOV-2001
 DEFINITION Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
 ACCESSION AF326353
 VERSION AF326353.1 GI:16797891
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 786)
 Holland, S.J., Liao, X.C., Mendenhall, M.K., Zhou, X., Pardo, J.,
 Chu, P., Spencer, C., Fu, A.C., Sheng, N., Yu, P., Pali, E., Nagin, A.,
 Shen, M., Yu, S., Chan, E., Wu, X., Li, C., Wolstschlaeger, M.,
 Aversa, G., Kolbinger, F., Bennett, M.K., Molineux, S., Luo, Y.,
 Payan, D.G., Mancebo, H.S.Y. and Wu, J.
 Functional Cloning of Src-like Adaptor Protein-2 (SLAP-2), a Novel
 Inhibitor of Antigen Receptor Signaling
 J. Exp. Med. 194 (9), 1263-1276 (2001)
 JOURNAL MEDLINE 21553259
 PUBMED 11696592
 REFERENCE 2 (bases 1 to 786)
 Holland, S.J., Mendenhall, M.K., Zhou, X., Spencer, C., Pardo, J.,
 Fu, A.C., Sheng, N., Shen, M., Liao, X.C., Luo, Y., Payan, D.G.,
 Mancebo, H.S.Y. and Wu, J.
 Direct Submission
 Submitted (05-DEC-2000) Rigol Pharmaceutical Inc., 240 East Grand
 Avenue, South San Francisco, CA 94080, USA
 Location/Qualifiers
 1..786
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 1..786
 /note="SLAP2; SH2 and SH3 domain-containing adaptor"
 /codon_start=1
 /product="Src-like adaptor protein-2"
 /protein_id="AAJ29204.1"
 /db_xref="GI:16797892"
 /translation="MGSLLPSSRRKSLPSPSSVQGGPVTMEERSKATVALGSP
 AGPAELSLRGEPLTIVSDGDMWTLSVSGREYNI PSYHAKVSHGWLVEGLSRE
 KAELLILPBNPGAFILIRSDQTRGYSLSVRLSPRSMIRIHYRIHCLDNGMLYI
 SPRLTPPSIQALVHYSHLADIDICLLKEPCVLRAPGLPKDIPLPVTQRPPLNWK
 ELDSLLFSEATGESLISRLRESLSFYISLNDVSLDDA"
 ORIGIN
 Query Match 100.0%; Score 786; DB 9; Length 786;

Best Local Similarity 100.0%; Pred. No. 8.5e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGAAGTCTGCGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAAGTCTCTGTC 60
DB 1 ATGGGAAGTCTGCGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAAGTCTCTGTC 60
QY 61 CAAAGCCAGGAGCTTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 61 CAAAGCCAGGAGCTTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 121 GGCAGATTTCCCGCAGAGTGGCCCGCGAGCTGTGCTGAGACTGCGGAGAGCCATTGACC 180
DB 121 GGCAGATTTCCCGCAGAGTGGCCCGCGAGCTGTGCTGAGACTGCGGAGAGCCATTGACC 180
QY 181 ATGCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
DB 181 ATGCTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
QY 241 AACATCCCCAGCGCTCCACGTCGCGCAAGAGTCTCCCATGAGTGGCTGATAGAGGCTGAGC 300
DB 241 AACATCCCCAGCGCTCCACGTCGCGCAAGAGTCTCCCATGAGTGGCTGATAGAGGCTGAGC 300
QY 301 AGGAGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 301 AGGAGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CGGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 CGGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TCCGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 TCCGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TCACCGGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGAGACCTTACTGAGCTGGCG 540
DB 481 TCACCGGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGAGACCTTACTGAGCTGGCG 540
QY 541 GATGACATCTGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGAGAGGCTGGCGCTGCTCT 600
DB 541 GATGACATCTGCGCTCACTTCCCTCTCACTCCAGGCGCTGTGAGAGGCTGGCGCTGCTCT 600
QY 601 GGCAGAGATATACCTCTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GGCAGAGATATACCTCTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 CTCCGAGAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780
DB 721 CTCCGAGAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 780
QY 781 GCTTAG 786
DB 781 GCTTAG 786

```

RESULT 5
AK025645 2415 bp mRNA linear PRI 13-SEP-2003
LOCUS Homo sapiens cDNA: FLJ21992 fls, clone HEP06554.
DEFINITION AK025645.1 GI:10438227
ACCESSION AK025645
VERSION o1igo capping; fls (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
Location/Qualifiers
1. 2415
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP06554"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_1ib="HEP"
/note="Cloning vector pME18PL3"
61. 846
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15201.1"
/db_xref="GI:10438228"
/translation="MGSLPRRSKSLPSLSVVOGGPYTMEASRKATVALGSP
AGSPARLSRLGRLPLTVSEDDGMWYLVSVSGREYNIPSVHAKYSHGLYGLSRE
KAEELILPENPGAFILRESQTRRSYSLSVLSRPSADRLRHRIRHCLDQNMXYI
SPRLTPPSLQALVDHYSELADICLLKEPCVLRAGPLPKDIPLEVTQRTPLNKK
ELDSILFSEATGEBESLSLSEGRBSLSEFYISINDENSLIDA"

ORIGIN
Query Match 100.0%; Score 786; DB 9; Length 2415;
Best Local Similarity 100.0%; Pred. No. 8.2e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGAAGTCTGCGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAAGTCTCTGTC 60
DB 61 ATGGGAAGTCTGCGCCAGCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAAGTCTCTGTC 120
QY 61 CAAAGCCAGGAGCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 121 CAAAGCCAGGAGCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 121 GGCAGATTTCCCGCAGAGTGGCCCGCGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAT 180
DB 181 GGCAGATTTCCCGCAGAGTGGCCCGCGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAT 240
QY 181 ATGCTCTGAGAGATGAT 240
DB 241 ATGCTCTGAGAGATGAT 300
QY 241 AACATCCCCAGCGCTCCACGTCGCGCAAGAGTCTCCCATGAGTGGCTGATAGAGGCTGAGC 300
DB 301 AACATCCCCAGCGCTCCACGTCGCGCAAGAGTCTCCCATGAGTGGCTGATAGAGGCTGAGC 360
QY 301 AGGAGAAAGCAG 360
DB 361 AGGAGAAAGCAG 420
QY 361 CGGAGAGAGCAG 420
DB 421 CGGAGAGAGCAG 480

```

QY 421 TCCTGGGACCGGATCGACACTACAGGATCCACTGCTTGAACATGCGCTGATCATC 480
DB 481 TCTTGGGACCGGATCGACACTACAGGATCCACTGCTTGAACATGCGCTGATCATC 540
QY 481 TCACCGCGGCTCACTTCCCTCACTCCAGCCCTGTGAGCATTAATCTAGAGCGGCG 540
DB 541 TCACCGCGGCTCACTTCCCTCACTCCAGCCCTGTGAGCATTAATCTAGAGCGGCG 600
QY 541 GATGACATCTGCTGCTTCACTCAAGAGAGCCCTGTGCTTCTGAGAGGCGGCTGCTCCT 600
DB 601 GATGACATCTGCTGCTTCACTCAAGAGAGCCCTGTGCTTCTGAGAGGCGGCTGCTCCT 660
QY 601 GGCAAGGATATACCCCTTACCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 661 GGCAAGGATATACCCCTTACCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 661 GACAGCTCCCTCTCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 721 GACAGCTCCCTCTCTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 721 CTCGGGAGTCCCTCACTTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB 781 CTCGGGAGTCCCTCACTTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 781 GCGCTAG 786
DB 841 GCGCTAG 846

```

RESULT 6
LOCUS BC042041
DEFINITION Homo sapiens Src-1-like-adaptor 2, transcript variant 1, mRNA (cDNA clone MGC:49845 IMAGE:4429896), complete cds.
ACCESSION BC042041
VERSION BC042041.1 GI:27469842
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Schenken, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., Moxham, K.J., Malek, U.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalobon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerker, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PubMed 12477932
REFERENCE Strausberg, R.
AUTHORS Submitted (23-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT

Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTM
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdelpaxll@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/liml.at: <http://image.llnl.gov>
Series: IPAK Plate: 88 Row: A Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 28416422.

FEATURES

source

1..2538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:49845 IMAGE:4429896"
/tissue_type="prostate, adenocarcinoma."
/clone_id="NIH_MGC_91"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..2538
/gene="SLA2"
/note="synonyms: SLAP-2, MGC49845, FLJ21992, SLAP2, C20orf156"
/db_xref="locusid:84174"
/db_xref="MIM:606577"
363..1148
/gene="SLA2"
/codon_start=1
/product="Src-1-like-adaptor 2, isoform a"
/protein_id="AA42041.1"
/db_xref="GI:27469843"
/db_xref="locusid:84174"
/db_xref="MIM:606577"
/translaton="MGSLLPSRRKSLPSLSVQGGPYTMEARSKATVALGSP AGGPAELSLRGEPLTIVSESDGMWTVLSVSGREYNI PSYHVKVSHMLYGLSRE KAEELLLPNPGGAFPIRESOTPRGYSLSVLRPSAMRIRHYHICLNDGMLTY SPRTFSLQALVDHSELDADICLKEPCVORAPLRGKDIPEVTYQRTPLNKK EDSLSLFSNAIGESLBSGLRESLSFYISLNDRAVSLDDA"

CDS

1 ATGGGAAGTCTGCCAGACAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTGTC 60
363 ATGGGAAGTCTGCCAGACAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTGTC 422

ORIGIN

Query Match 100.0%; Score 786; DB 9; Length 2538;
Best Local Similarity 100.0%; Pred. No. 8 2e-195;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCCAGACAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTGTC 60
DB 363 ATGGGAAGTCTGCCAGACAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTCTCTGTC 422
QY 61 CAAGGCGAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 423 CAAGGCGAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 482
QY 121 GAGAGTTTCCCGGAGGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 180
DB 483 GAGAGTTTCCCGGAGGTGGCCCGGCGGAGCTGTGCTGAGACTCGGGAGCCATTGACC 542
QY 181 ATGCTCTGAGAGATGAGACTGAGTGGAGCGGTGCTGTCTGAAGTCTGAGCGAGAGATAT 240
DB 543 ATGCTCTGAGAGATGAGACTGAGTGGAGCGGTGCTGTCTGAAGTCTGAGCGAGAGATAT 602
QY 241 AACATCCCAAGCTTCACGTCGAGCAAGTCTCCCATGAGGAGGCTGATGAGGCGCTAGC 300
DB 603 AACATCCCAAGCTTCACGTCGAGCAAGTCTCCCATGAGGAGGCTGATGAGGCGCTAGC 662
QY 301 AGGAGAGAAAGCAGAGAGAACTGCTGTTTACTGAGAAACCTTGAGGAGGCGCTTCTCATC 360

Db 663 AGGAGAAACGAGGAACTGCTTTGTTTACTCGGAAACCTTGAGGGGCTTTCCTCATT 722
QY 361 CGGAGAGCCAGACCGAGAGGCTTCTACTCTCTAGTCGCTTACGCGGCTTGA 420
Db 723 CGGAGAGCCAGACCGAGAGGCTTCTACTCTCTAGTCGCTTACGCGGCTTGA 782
QY 421 TCCTGGGACCGAGATCGAGACTACAGAGATCCACTGCTTGACAAATGGCTGGCTGATC 480
Db 783 TCCTGGGACCGAGATCGAGACTACAGAGATCCACTGCTTGACAAATGGCTGGCTGATC 842
QY 481 TCACCGGCTTCACTTCCCTCTACTCAAGGCTTGTGAGCCATTACTGAGCTGGCG 540
Db 843 TCACCGGCTTCACTTCCCTCTACTCAAGGCTTGTGAGCCATTACTGAGCTGGCG 902
QY 541 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTCTGAGAGGGCTGGCCCTCCCT 600
Db 903 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTCTGAGAGGGCTGGCCCTCCCT 962
QY 601 GCGAAGATATACCTCTACTGAGCTGAGAGGACACCACTCACTGAGAAAGAGCTG 660
Db 963 GCGAAGATATACCTCTACTGAGCTGAGAGGACACCACTCACTGAGAAAGAGCTG 1022
QY 661 GACAGCTCCCTCTCTGTTTCTGAACTGCCACAGAGGAGGTCTTCTCACTGAGAGGT 720
Db 1023 GACAGCTCCCTCTCTGTTTCTGAACTGCCACAGAGGAGGTCTTCTCACTGAGAGGT 1082
QY 721 CTCGGGAGATCCCTCACTTCTACATAGGCTGAATGACAGAGCTGCTTCTTGATGAT 780
Db 1083 CTCGGGAGATCCCTCACTTCTACATAGGCTGAATGACAGAGCTGCTTCTTGATGAT 1142
QY 781 GCCTAG 786
Db 1143 GCCTAG 1148

RESULT 7
AX452880 2567 bp DNA linear PAT 06-JUL-2002
LOCUS AX452880
DEFINITION Sequence 1 from Patent WO0242457.
ACCESSION AX452880
VERSION AX452880.1 GI:21712520
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villari, J.J. and Kanter, S.B.
TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific expression
JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
FEATURES
source location/Qualifiers
1..2567
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 786; DB 6; Length 2567;
Best Local Similarity 100.0%; Pred. No. 8.2e-195; Indels 0; Gaps 0;
Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAAGTCTCTGTC 60
Db 415 ATGGGAATCTGCCAGCAGAGAAATCTCTGCCAGCCCAAGCTTGAAGTCTCTGTC 474
QY 61 CAAGGCCAGGACTCTGTGACATGAGAGCAGAGAGAGCCACAGCCCTGGCCCTG 120
Db 475 CAAGGCCAGGACTCTGTGACATGAGAGCAGAGAGAGCCACAGCCCTGGCCCTG 534

QY 121 GGCAGTTTCCCGGAGAGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCTATTGAC 180
Db 535 GGCAGTTTCCCGGAGAGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCTATTGAC 594
QY 181 ATGCTCTGAGAGATGAGACTGTGTGAGAGAGGTCTGTGAAGTCTCAGGAGAGAT 240
Db 595 ATGCTCTGAGAGATGAGACTGTGTGAGAGAGGTCTGTGAAGTCTCAGGAGAGAT 654
QY 241 AACATCCCGAGGATCCAGCTGGCCAAAGTCTCCCATGGGTGAGCTGATAGAGGCTGAGC 300
Db 655 AACATCCCGAGGATCCAGCTGGCCAAAGTCTCCCATGGGTGAGCTGATAGAGGCTGAGC 714
QY 301 AGGAGAAACGAGAGAACTGCTGTTGTTTAACTGAGAACCTTGGAGGGCTTCTCATT 360
Db 715 AGGAGAAACGAGAGAACTGCTGTTGTTTAACTGAGAACCTTGGAGGGCTTCTCATT 774
QY 361 CGGAGAGCCAGACCGAGAGAGGCTTACTCTCTGTCAGTGGCCCTCAGCGGCTTGA 420
Db 775 CGGAGAGCCAGACCGAGAGAGGCTTACTCTCTGTCAGTGGCCCTCAGCGGCTTGA 834
QY 421 TCCTGGGACCGAGATCGAGACTACAGAGATCCACTGCTTGACAAATGGCTGATCATT 480
Db 835 TCCTGGGACCGAGATCGAGACTACAGAGATCCACTGCTTGACAAATGGCTGATCATT 894
QY 481 TCACCGGCTTCACTTCCCTCTACTCAAGGCTTGTGAGCCATTACTTCTGAGCTGGCG 540
Db 895 TCACCGGCTTCACTTCCCTCTACTCAAGGCTTGTGAGCCATTACTTCTGAGCTGGCG 954
QY 541 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTCTGAGAGGGCTGGCCCTCCCT 600
Db 955 GATGACATCTGCTGCTTACTCAAGAGCCCTGTGCTCTGAGAGGGCTGGCCCTCCCT 1014
QY 601 GCGAAGATATACCTCTACTGAGCTGAGAGGACACCACTCACTGAGAAAGAGCTG 660
Db 1015 GCGAAGATATACCTCTACTGAGCTGAGAGGACACCACTCACTGAGAAAGAGCTG 1074
QY 661 GACAGCTCCCTCTCTGTTTCTGAACTGCCACAGAGGAGAGGTCTTCTCACTGAGAGGT 720
Db 1075 GACAGCTCCCTCTCTGTTTCTGAACTGCCACAGAGGAGAGGTCTTCTCACTGAGAGGT 1134
QY 721 CTCGGGAGATCCCTCACTTCTACATAGGCTGAATGACAGAGGCTGCTTCTTGATGAT 780
Db 1135 CTCGGGAGATCCCTCACTTCTACATAGGCTGAATGACAGAGGCTGCTTCTTGATGAT 1194
QY 781 GCCTAG 786
Db 1195 GCCTAG 1200

RESULT 8
AX780857 2788 bp DNA linear PAT 14-JUL-2003
LOCUS AX780857
DEFINITION Sequence 3014 from Patent WO03039443.
ACCESSION AX780857
VERSION AX780857.1 GI:32697851
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Haeflisch, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S., Dugas, M., Ellis, R., Broers, B. and Wergenthaler, S.
TITLE Novel genetic markers for leukemias
JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;
FEATURES
source location/Qualifiers
1..2788
/organism="Homo sapiens"
/mol_type="unassigned DNA"

REFERENCE
AUTHORS Haeflisch, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S., Dugas, M., Ellis, R., Broers, B. and Wergenthaler, S.
TITLE Novel genetic markers for leukemias
JOURNAL Patent: WO 03039443-A 3014 15-MAY-2003;
FEATURES
source location/Qualifiers
1..2788
/organism="Homo sapiens"
/mol_type="unassigned DNA"

ORIGIN /db_xref="taxon:9606"

Query Match 99.9%; Score 785; DB 6; Length 2788;
Best Local Similarity 99.9%; Pred. No. 1.5e-194;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGCTGTGCCCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTG 60
Db 387 ATGGGAAGCTGTGCCCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTG 446

QY 61 CAAGGCCAAGGACCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 447 CAAGGCCAAGGACCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 506

QY 121 GGCAAGTTTCCCGGCAAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 180
Db 507 GGCAAGTTTCCCGGCAAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 566

QY 181 ATCGTCTGAT 240
Db 567 ATCGTCTGAT 626

QY 241 AACATCCCCAGCGTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 627 AACATCCCCAGCGTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 686

QY 301 AGGGAAG 360
Db 687 AGGGAAG 746

QY 361 CGGAG 420
Db 747 CGGAG 806

QY 421 TCTGGGACCGGATCAGACATCAGAGATCCAGTCCCTTGAACATGAGTGGCTGTATATC 480
Db 807 TCTGGGACCGGATCAGACATCAGAGATCCAGTCCCTTGAACATGAGTGGCTGTATATC 866

QY 481 TCACCGGCGCTCAGCTTCCCTCACTCCAGAGCCCTGTGTGACATTAATCTGAGCTGAGC 540
Db 867 TCACCGGCGCTCAGCTTCCCTCACTCCAGAGCCCTGTGTGACATTAATCTGAGCTGAGC 926

QY 541 GATGACATCTGCTGCTCACTCAAGAGAGCCCTGTGTGCTGAGAGAGAGAGAGAGAG 600
Db 927 GATGACATCTGCTGCTCACTCAAGAGAGCCCTGTGTGCTGAGAGAGAGAGAGAGAG 986

QY 601 GGCAAGGATATACCCCTTACCTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 987 GGCAAGGATATACCCCTTACCTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046

QY 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 1047 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1106

QY 721 CTCGGGAGTCCCTCAGCTTCTATAGAGCTGAAATGACAGAGAGAGAGAGAGAGAGAT 780
Db 1107 CTCGGGAGTCCCTCAGCTTCTATAGAGCTGAAATGACAGAGAGAGAGAGAGAGAT 1166

QY 781 GCCTAG 786
Db 1167 GCCTAG 1172

RESULT 9
AX443133 1183 bp DNA linear PAT 02-JUL-2002
LOCUS AX443133
DEFINITION Sequence 74 from Patent WO0216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690555
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1183
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 99.8%; Score 784.4; DB 6; Length 1183;
Best Local Similarity 99.9%; Pred. No. 2.2e-194;
Matches 785; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAAGCTGTGCCCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTG 60
Db 398 ATGGGAAGCTGTGCCCCAGAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTTCTCTGTG 457

QY 61 CAAGGCCAAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 458 CAAGGCCAAGGACCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517

QY 121 GGCAAGTTTCCCGGCAAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 180
Db 518 GGCAAGTTTCCCGGCAAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCCATTGACC 577

QY 181 ATCGTCTGAT 240
Db 578 ATCGTCTGAT 637

QY 241 AACATCCCCAGCGTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 638 AACATCCCCAGCGTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697

QY 301 AGGGAAG 360
Db 698 AGGGAAG 757

QY 361 CGGAG 420
Db 758 CGGAG 817

QY 421 TCTGGGACCGGATCAGACATCAGAGATCCAGTCCCTTGAACATGAGTGGCTGTATATC 480
Db 818 TCTGGGACCGGATCAGACATCAGAGATCCAGTCCCTTGAACATGAGTGGCTGTATATC 877

QY 481 TCACCGGCGCTCAGCTTCCCTCACTCCAGAGCCCTGTGTGACATTAATCTGAGCTGAGC 540
Db 878 TCACCGGCGCTCAGCTTCCCTCACTCCAGAGCCCTGTGTGACATTAATCTGAGCTGAGC 937

QY 541 GATGACATCTGCTGCTCACTCAAGAGAGCCCTGTGTGCTGAGAGAGAGAGAGAGAG 600
Db 938 GATGACATCTGCTGCTCACTCAAGAGAGCCCTGTGTGCTGAGAGAGAGAGAGAGAG 997

QY 601 GGCAAGGATATACCCCTTACCTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 998 GGCAAGGATATACCCCTTACCTGTGACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1057

QY 661 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 1058 GACAGCTCCCTCTCTGTTTCTGAAAGCTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1117

QY 721 CTCGGGAGTCCCTCAGCTTCTATAGAGCTGAAATGACAGAGAGAGAGAGAGAGAT 780
Db 1118 CTCGGGAGTCCCTCAGCTTCTATAGAGCTGAAATGACAGAGAGAGAGAGAGAT 1177

QY 781 GCCTAG 786


```

Db      421 TCCTGGAGCCGGATGACAGACTACAGACTGCTGCTTGAAGTGGCTGCTGATC 480
Qy      481 TCACCGGCGCTCACTTCCCTCCTACTCCAGGCGCTGGTGGACATTACTGAGCTGAG 540
Db      481 TCACCGGCGCTCACTTCCCTCCTACTCCAGGCGCTGGTGGACATTACTCT----- 531
Qy      541 GATGACATGTGCTGCTACTCAAGAGCCCTGTGTCTGACAGAGGCTGGCCGCTCCT 600
Db      532 -----GAGGCTGGCCGCTCCT 550
Qy      601 GGCAAGATATATACCCCTACTGCTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 660
Db      551 GGCAAGATATATACCCCTACTGCTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 610
Qy      661 GACAGCTCCTCTCTGTTTCTGAGCTGCGACAGAGAGAGTCTCTTCTCAGTGAAGGT 720
Db      611 GACAGCTCCTCTCTGTTTCTGAGCTGCGACAGAGAGAGTCTCTTCTCAGTGAAGGT 670
Qy      721 CTCGGAGAGTCCCTGAGCTTCTACTCATCATCAGCCCTGAATGA-CGAGGCTGTCTCTTGGATGA 779
Db      671 CTCGGAGAGTCCCTGAGCTTCTACTCATCATCAGCCCTGAATGACGAGGCTGTCTCTTGGATGA 730
Qy      780 TGCCCTAG 786
Db      731 TGCCCTAG 737

```

```

RESULT 12
LOCUS   AF290986              737 bp  mRNA  linear  PRI 21-JAN-2003
DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
complete cds; alternatively spliced.
ACCESSION AF290986
VERSION   AF290986.1 GI:17351922
KEYWORDS
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     1 (bases 1 to 737)
JOURNAL   Loreto, M.P. and McGlade, C.J.
MEDLINE   Cloning and characterization of human Src-like adaptor protein 2
22415750
JOURNAL   Oncogene 22 (2), 266-273 (2003)
REFERENCE 2 (bases 1 to 737)
AUTHORS  Loreto, M.P. and McGlade, C.J.
JOURNAL   Direct Submission
TITLE     Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
Source
1. 737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="20"
/tissue_type="thymus"
1. 633
/feature="SLAP-2-v"
/codon_start=1
/product="Src-like adaptor protein-2 splice isoform"
/protein_id="P481819.1"
/db_xref="GI:17351923"
/transcript="M581P5RKLPSLPSLSSSSVQGGPYTMEAREKATVALGSP
AGGPAELSLRLEPLTIVSEDDPMWTVLSVSGREYVTPVHAKVSHGMLYGLSP
KAEELLLPENGAPLIRSGOTRSGYSVLSVPSASDRLRHRYIHLIDNGMLYI
SPRLTFSGLALVDHYSBGMPPAQGTTPCTCDABDTTQLERAGQLPVE"

```

```

CDS
1. 633
/feature="SLAP-2-v"
/codon_start=1
/product="Src-like adaptor protein-2 splice isoform"
/protein_id="P481819.1"
/db_xref="GI:17351923"
/transcript="M581P5RKLPSLPSLSSSSVQGGPYTMEAREKATVALGSP
AGGPAELSLRLEPLTIVSEDDPMWTVLSVSGREYVTPVHAKVSHGMLYGLSP
KAEELLLPENGAPLIRSGOTRSGYSVLSVPSASDRLRHRYIHLIDNGMLYI
SPRLTFSGLALVDHYSBGMPPAQGTTPCTCDABDTTQLERAGQLPVE"
ORIGIN
Query Match 84.4%; Score 663.4; DB 9; Length 737;
Best Local Similarity 93.4%; Pred. No. 1,1e-162;

```

```

Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;
Qy      1 ATGGAAAGTCTGCCAGCAGAGAAATCTGCCAAGCCCAAGCTTGAATCTCTGTC 60
Db      1 ATGGAAAGTCTGCCAGCAGAGAAATCTGCCAAGCCCAAGCTTGAATCTCTGTC 60
Qy      61 CAAGCCAGAGGACCTGTGACATGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
Db      61 CAAGCCAGAGGACCTGTGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 120
Qy      121 GGCATTTTCCCGGACAGGTGGCCCGCCGAGCTGTGCTGAGACTGGGAGGCAATTGACC 180
Db      121 GGCATTTTCCCGGACAGGTGGCCCGCCGAGCTGTGCTGAGACTGGGAGGCAATTGACC 180
Qy      181 ATGCTCTGAGAGATGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
Db      181 ATGCTCTGAGAGATGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
Qy      241 AACATCCCGACGCTCAAGTGGCCAAAGTCTCCATGGAGTGGCTGTATAGAGGCTGAGC 300
Db      241 AACATCCCGACGCTCAAGTGGCCAAAGTCTCCATGGAGTGGCTGTATAGAGGCTGAGC 300
Qy      301 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360
Db      301 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 360
Qy      361 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 420
Db      361 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 420
Qy      421 TCCTGGAGCCGAGATCAGACATCAAGATCCACTGCTTGAACAATGAGTGGCTGTACATC 480
Db      421 TCCTGGAGCCGAGATCAGACATCAAGATCCACTGCTTGAACAATGAGTGGCTGTACATC 480
Qy      481 TCACCGGCGCTCAGCTTCTCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 540
Db      481 TCACCGGCGCTCAGCTTCTCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 540
Qy      541 GATGACATGTGCTGCTACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
Db      532 -----GAGGCTGGCCGCTCCT 550
Qy      601 GGCAAGATATATACCCCTACTGCTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 660
Db      551 GGCAAGATATATACCCCTACTGCTGACTGTGACAGAGACCACTCAACTGAAAGAGCTG 610
Qy      661 GACAGCTCCTCTCTGTTTCTGAGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAT 720
Db      611 GACAGCTCCTCTCTGTTTCTGAGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAT 670
Qy      721 CTCGGAGAGTCCCTGAGCTTCTACTCATCATCAGCCCTGAATGA-CGAGGCTGTCTCTTGGATGA 779
Db      671 CTCGGAGAGTCCCTGAGCTTCTACTCATCATCAGCCCTGAATGACGAGGCTGTCTCTTGGATGA 730
Qy      780 TGCCCTAG 786
Db      731 TGCCCTAG 737

```

```

RESULT 13
LOCUS   BC052655              1384 bp  mRNA  linear  ROD 04-NOV-2003
DEFINITION Mus musculus Src-like-adaptor 2, mRNA (cDNA clone MGC:60811
IMAGE:30040401), complete cds.
ACCESSION BC052655
VERSION   BC052655.1 GI:30851667
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1384)
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

```


ORIGIN

Query Match 68.4%; Score 537.8; DB 10; Length 1348;
Best Local Similarity 81.9%; Pred. No. 8e-130;
Matches 646; Conservative 0; Mismatches 137; Indels 6; Gaps 2;

QY 1 ATGGGAAGTCTGCCCGCAGAGAAAATCTGCGCAAGCCCAAGCTTGAGTTCCTCTGTC 60
DB 282 ATGGGAAGTTCCTGCGAGAGGAAAAC---CTCCAGCCCGACCCCGAGCTCCTTGGT 338
QY 61 CAAGGCCAGGGAAGCTGTGACCATGGAAGAGAGAAAGCAAGGCCACAGCCGTGGCCCTG 120
DB 339 CCAGACCAAGAAACCGGTGTCCATGCAACCAAGAAAGACAGAGGTCAAGCTGTGGCCCTG 398
QY 121 GGCAGTTTCCCGCAGAGTGGCCCGCGCAGCTGTGCTGAGACTCGGGGAGCCATTGACC 180
DB 399 GGCAGTTTCCCGCAGAGTGAACAGGCCAAGCTATCTGTGAGACTCGGGGAGCCGTGACC 458
QY 181 ATGCTCTGAGAGTGAAGCTGTGAGCGGTGCTGTGTAAGTCTCAGGCAGAGAGTAT 240
DB 459 ATCATCTTGAGAGTGAAGATTGTGGAACATCCAGTCCGAAAGTCTCAGGCAGAGATAC 518
QY 241 AACATCCCAAGCGTCCACGTGCGCAAAAGTCTCCCATGGGTGCTGTATGAGGGCTGAGC 300
DB 519 CACATGCCCGAGTGTATGTGGCTAAAGTCCCGACGGGTGGCTGTACAGAGGGCTGAGC 578
QY 301 AGGGAAGAAAGCAGAGAACTGCTGTGTATCCTGGGAACCTTGAGGGGCTTCTCATC 360
DB 579 CGGAGAAAGCCAGAGAACTACTCTGTATCCTGGGAACCCCGAGGGGCTTCTCATC 638
QY 361 CGGAGAGCCAGACCAAGAGAGGCTTTACTCTGTGAGTCCGCTCAGCCGCGCTGCA 420
DB 639 CGGAGAGCCAGACCAAGAGAGGCTGTATTCCTGTCCGTCCGACTCAGCCGCGCTGCA 698
QY 421 TCCTGGAGCCGATCAGACACTACAGATCCACTGCTTGAACATGCTGCTGTATC 480
DB 699 TCTTGGGACCGGATCAGACACTACAGATACAGGTTCCTTGAACATGCTGCTGTATC 758
QY 481 TCACCGCGCTCACTTCCCTCTCACTCCAGGCCCTGGTGGACCATTACTCTGAGCTGCG 540
DB 759 TCACCTGCGCTCACTTCCCTCTCACTCCAGGCCCTGGTGGACCATTACTCTGAGCTGCA 818
QY 541 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTCTGCGAGAGGCTGGCCGCTCCCT 600
DB 819 GATGCAATCTGCTGCTCCCTCAGGAGCCGTGTCTCTGCGAGAGCTTGGGCCACTACCT 878
QY 601 GGCAAGGATATACCCCTAAGCTGTGACAGAGACCACTCAACTGGAAGAGCTG 660
DB 879 GGCAAGGATATACCTCAAGCTGTGACAGAGCAATCATATCACTAAATGGAAGAGCTG 938
QY 661 GACAGCTCCCTCTGTTTCTGAG---CTGCCAAGGGGAGAGTCTCTTCTCAGTGAG 717
DB 939 GACGCGAGCCCTCTGTTTCTGGAAGCACTCGAGTGGGAGGCACTCTCTCAGTGAG 998
QY 718 GGTCTCCGGAGTCCCTCAGCTTCTACATCAGCCGTAATGACGAGGCTGTCTCTTTGAT 777
DB 999 GGGCTCCGAGAGTCCCTCAGTTCCTACATCAGCCGCTGAGGACCCCTTGATGATGCT 1058
QY 778 GATGCTTAG 786
DB 1059 TAGCCCTGG 1067

Search completed: November 16, 2004, 22:36:22
Job time : 3774 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2004, 23:32:28 ; Search time 4546 Seconds
(without alignment)

2715.049 Million cell updates/sec

Title: US-10-043-649-2

Perfect score: 1351

Sequence: 1 MGSLLPSSRRKSLPSPSSSV.....RELSFTSLNDEAVSLDDA 261

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10043649/runat.16112004.060536.28719/app.query.fasta.1.455
-DB=genemb1 -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOADCL=0 -LOADEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=200000000
-USER=US10043649@cgn.1_1371 @runat.16112004.060536.28719 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmb1.*
1: gb_da.*
2: gb_hc9.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_pl.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1351	100.0	786	6	AX511153
2	1351	100.0	786	6	AX572845
3	1351	100.0	786	9	AF290985
4	1351	100.0	786	9	AF326353

5	1351	100.0	2415	9	AK025645	AK025645 Homo sapi
6	1351	100.0	2538	9	BC042041	BC042041 Homo sapi
7	1351	100.0	2567	6	AX452880	AX452880 Sequence
8	1347	99.7	1183	6	AX443133	AX443133 Sequence
9	1347	99.7	1183	6	AX443135	AX443135 Sequence
10	1345	99.6	2788	6	AX780857	AX780857 Sequence
11	1200.5	88.9	737	6	AX511155	AX511155 Sequence
12	1200.5	88.9	737	6	AF290986	AF290986 Homo sapi
13	1032	76.4	777	6	AX511151	AX511151 Sequence
14	1032	76.4	1348	6	AX511150	AX511150 Sequence
15	1032	76.4	1348	10	AF287467	AF287467 Mus muscu
16	1032	76.4	1384	10	BC052655	BC052655 Mus muscu
17	1023	75.7	1321	10	AF434990	AF434990 Mus muscu
18	770	57.0	441	6	C0724864	C0724864 Sequence
19	647	47.9	864	6	C0414231	C0414231 Sequence
20	587	43.4	1571	5	BC056035	BC056035 Xenopus l
21	505	37.4	1631	10	AY217759	AY217759 Rattus no
22	496	36.7	921	10	MMU29056	MMU29056 Mus muscu
23	496	36.7	2648	10	BC032922	BC032922 Mus muscu
24	496	36.7	2773	10	MMU11777	MMU11777 Mus muscu
25	495.5	36.7	841	10	AY079449	AY079449 Mus muscu
26	495.5	36.7	841	10	AY079450	AY079450 Mus muscu
27	491	36.3	1076	9	HSU30473	HSU30473 Homo sapien
28	491	36.3	1870	9	BC007042	BC007042 Homo sapi
29	491	36.3	2021	9	HSU44403	HSU44403 Human Src-1
30	491	36.3	2109	6	AX428893	AX428893 Sequence
31	491	36.3	2665	6	AX333017	AX333017 Sequence
32	491	36.3	2665	9	DB9077	DB9077 Homo sapien
33	491	36.3	2670	6	C0725810	C0725810 Sequence
34	491	36.3	3090	9	HSM808530	HSM808530 Homo sapi
35	485.5	35.9	831	9	CR536537	CR536537 Homo sapi
36	482	35.7	145068	9	HSU797781	HSU797781 Homo DNA
37	482	35.7	145833	2	AC026539	AC026539 Homo sapi
38	471.5	34.9	849	5	AY278230	AY278230 Gallus ga
39	384.5	28.5	2703	10	MUSLINB	MUSLINB Rattus norv
40	379.5	28.1	1538	10	RATLYNBYR	RATLYNBYR Rattus norv
41	379.5	28.1	1538	10	RATLYNBYR	RATLYNBYR Rattus norv
42	379.5	28.1	1538	10	RATLYNBYR	RATLYNBYR Rattus norv
43	374.5	27.7	1361	6	C0729540	C0729540 Sequence
44	374.5	27.7	1539	12	AY335782	AY335782 Synthetic
45	374.5	27.7	1613	9	BC075001	BC075001 Homo sapi

ALIGNMENTS

RESULT 1	AX511153	AX511153	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002
LOCUS	AX511153	AX511153	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002
DEFINITION	AX511153	AX511153	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002
ACCESSION	AX511153	AX511153	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002
VERSION	AX511153.1	GI:23392046	Sequence 4 from Patent WO0242452.	786 bp	DNA	linear	PAT 27-SEP-2002
KEYWORDS							
SOURCE							
ORGANISM							
	Homo sapiens (human)						
	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE							
1	Moglad, J.C. and Loreto, M.P.						
AUTHORS							
TITLE	Adapter gene						
JOURNAL	Patent: WO 0242452-A 4 30-MAY-2002;						
	The Hospital for Sick Children (CA)						
FEATURES							
source							
	1..786						
	/organism="Homo sapiens"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:9606"						

ORIGIN

Alignment Scores: 1.37e-101 Length: 786
Pred. No.: 1351.00 Matches: 261
Score: 1351.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00%

Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-043-649-2 (1-261) x AX511153 (1-786)			
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
DB	1	ATGGGAAGTCTGGCCAGAGAGAAATCTCTCCAGAACCCAGCTTGAATCTCTCTGTC	60
QY	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
DB	61	CAAGGCGAGGAGCCTGTGACCATGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	120
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlyProLeuThr	60
DB	121	GGGAGTTTCCGGAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGGAGGCAATTGACC	180
QY	61	IleValSerGluAspGlyAspTyrPheThrValLeuSerGluValSerGlyArgGlyTyr	80
DB	181	ATCGTCTCTGAGGATGAGACTGGTGGACGGTGTCTGAGAGTCTCAGGCGAGAGTAT	240
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTyrPheLeuTyrGlyLeuSer	100
DB	241	AACATCCCGACGCTCCACGCTGGCCAAAGTCTCCATGAGTGGCTGTATGAGGGCTGAGC	300
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120
DB	301	AGGAGAGAAAGAGAGAACTGCTTGTACTTGGAAACCTGGAGGGGCTTCTCTATC	360
QY	121	ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	361	CGGAGAGCCAGACAGAGAGAGGCTTACTCTCTGTCAGTCCGCTCAGCGCCCTGCA	420
QY	141	SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheTyrIle	160
DB	421	TCCGGAGACCGAGTCAACACACATCAAGATCCATGCTGACATGGCTGAGCTTACATC	480
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
DB	481	TCAACCGGCTCCTCACTTCCCTCACTCCAGGCCCTGTGGACCATTTACTGAGACTGGG	540
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
DB	541	GATGACATCTGCTGCTACTCAAGAGAGCCGTGTCTGCTGAGAGGGCTGCTCCCT	600
QY	201	GlyTyrAspIlePheProLeuProValThrValGlnArgThrProLeuAsnTyrPheGluLeu	220
DB	601	GGCAGAGATATATCCCTCACTGACTGTGCAGAGAGACCACTCAACTGAGAAAGAGCTG	660
QY	221	AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGlyGly	240
DB	661	GACAGCTCCCTCTGTTTCTGAACTGCCACAGGGAGAGAGTCTCTTCTCAAGTGAAGGT	720
QY	241	LeuArgIleSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp	260
DB	721	CTCCGGAGAGTCCCTCAGCTTCAATCACTGAGCTGAGATGACAGAGGCTGTCTTTGGATGAT	780
QY	261	Ala 261	
DB	781	GCC 783	

RESULT 2
 AX572845 786 bp DNA linear PAT 29-NOV-2002
 LOCUS Sequence 1 from Patent WO02055707.
 DEFINITION AX572845
 ACCESSION AX572845
 VERSION AX572845.1 GI:26004935
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS		Holland, S.J., Mendenhall, M.K., Pardo, J., Spencer, C., Fu, A.C., Luo, Y., Payan, D.G., Mancebo, H.S., Wu, J., Zhou, X., Shen, M., Liao, X.C. and Sheng, N.	
TITLE		Cloning of an inhibitor of antigen-receptor signaling by a retroviral-based functional screen	
JOURNAL		Patent: WO 02055707-A 1 18-JUL-2002; Rigel Pharmaceuticals, Inc. (US)	
FEATURES		Location/Qualifiers	
source		1..786	
CDS		/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 1..786 /note="unassigned protein product" /codon_start=1 /protein_id="CAD58542.1" /db_xref="GI:26004936" /translation="MSGSLPSRKSLPSLSVGGOGPYTWMEKRSKATAVAGSP AGGPAELSLRLSEPLTVSDDMTVLSVSRREYNTPSVYAKSHKLVGLSRE KAEELLLPQNPQAFILRSQTRGSYSLSVLSPPASMDRLRHRIHCLDGMAT SPLTPPSQALVDHYSELADICCLKEPCVLRQAGPLKDIPLPVQRPPLWKK ELDSSLFSEATGSESLSEGLRESLSFYSLNDAVSLDDA"	
ORIGIN			
Alignment Scores:			
Pred. No.:		1.37e-101	
Score:		1351.00	
Percent Similarity:		100.00%	
Best Local Similarity:		100.00%	
Query Match:		100.00%	
DB:		6	
US-10-043-649-2 (1-261) x AX572845 (1-786)			
QY	1	MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal	20
DB	1	ATGGGAAGTCTGGCCAGAGAGAAATCTCTCCAGAACCCAGCTTGAATCTCTCTGTC	60
QY	21	GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu	40
DB	61	CAAGGCGAGGAGCCTGTGACCATGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	120
QY	41	GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGlyProLeuThr	60
DB	121	GGCAGTTTCCCGAGGTGGCCCGCCGAGCTGTCTGAGACTCGGGGAGACCATTTGACC	180
QY	61	IleValSerGluAspGlyAspTyrPheThrValLeuSerGluValSerGlyArgGlyTyr	80
DB	181	ATCGTCTCTGAGGATGAGACTGGTGGACGGTGTCTGAGAGTCTCAGGCGAGAGATAT	240
QY	81	AsnIleProSerValHisValAlaLysValSerHisGlyTyrPheLeuTyrGlyLeuSer	100
DB	241	AACATCCCGACGCTCCACGCTGGCCAAAGTCTCCATGAGTGGCTGTATGAGGGCTGAGC	300
QY	101	ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle	120
DB	301	AGGAGAGAAAGAGAGAACTGCTTGTACTTGGAAACCTGGAGGGGCTTCTCTATC	360
QY	121	ArgIleSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla	140
DB	361	CGGAGAGCCAGACAGAGAGGCTTACTCTGTACATCCGCTCAGCGCCCTGCA	420
QY	141	SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheTyrIle	160
DB	421	TCCGGAGACCGAGTCAACACACATCAAGATCCATGCTGACATGGCTGAGCTTACATC	480
QY	161	SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla	180
DB	481	TCAACCGGCTCCTCACTTCCCTCACTCCAGGCCCTGTGGACCATTTACTGAGCTGGG	540
QY	181	AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro	200
DB	541	GATGACATCTGCTGCTACTCAAGAGAGCCCTGTGTCTGAGAGAGGGCTGCGCCCTCT	600

```

QY      201  GILYASPIIleProleuProValThrValGlnAqThProleuAsnTrpIysGluLeu 220
        |||
Db      601  GGCAAGATATACCCCTACTGCTGAGTGCAGAGACACCACTCAACCTGGAAGAGCTG 660
QY      221  AASPserLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
        |||
Db      661  GACAGCTCCCTCTCTTTCTGAGTGCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT 720
QY      241  LeuArgGluSerLeuSerPheTyrlleSerLeuAsnAspGluAlaValSerLeuAsp 260
        |||
Db      721  CTCGGAGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGCTGTCTCTTGATAT 780
QY      261  Ala 261
        |||
Db      781  GCC 783

RESULT 3
AF290985      786 bp  mRNA  linear  PRI 21-JAN-2003
LOCUS        Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION   AF290985
ACCESSION    AF290985.1 GI:17351920
VERSION      AF290985.1
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 786)
AUTHORS      Loreto,M.P. and McGlade,C.J.
TITLE        Cloning and characterization of human Src-like adaptor protein 2
              and a novel splice isoform, SLAP-2-v
JOURNAL      Oncogene 22 (2), 266-273 (2003)
MEDLINE      22415750
PUBMED       12527895
REFERENCE    2 (bases 1 to 786)
AUTHORS      Loreto,M.P. and McGlade,C.J.
TITLE        Direct Submission
JOURNAL      Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
              Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
FEATURES
             source
             1..786
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /chromosome="20"
               /tissue_type="thymus"
               1..786
               /note="SLAP-2"
               /codon_start=1
               /product="Src-like adaptor protein-2"
               /protein_id="AAL38197.1"
               /db_xref="GI:17351921"
               /translation="MGSLSPRKSLPSPLSSSVGGOGPYTMEARSKATAVALGSP
               AGGAELSLRLGEPLTIVSEDDGWWTALSTFGSGEKNVPSHVAKUSHWIYEGLSRE
               KAERLLPGRPGAFILRESQTRGYSISVRLSPASWDRIRHITHICIDNEMLYI
               SPRTFSPQLAVVHSELDADICLKEPVLQRAPLPKDKLPLVTVQRPFLMWK
               ELDSLLFSEATRESLISBGLRESLSPYISINDEAVSIDDA"
ORIGIN
Alignment Scores:
Pred. No.:      1,37e-101      Length:      786
Score:          1351.00      Matches:      261
Percent Similarity: 100.00%      Mismatches: 0
Best Local Similarity: 100.00%      Indels:      0
Query Match:    100.00%      Gaps:        0
US-10-043-649-2 (1-261) x AF290985 (1-786)
QY      1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
        |||
Db      1  ATGGAGAGTGTGCCACAGAGAAATCTCTGCGCAAGCCCAAGCTTGAGTTCTCTGTC 60

```

```

QY      21  GlnGlyGlnGlyProValThrMetGluAlaGluArgSerValaThrAlaValAlaLeu 40
        |||
Db      61  CAAGCCAGAGGAGCCTGTGATACATGAGACAGAGAGACAGACAGCCCTGGCCCTG 120
QY      41  GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
        |||
Db      121  GGCAGCTTCCCGGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGGCATTGACC 180
QY      61  IleValSerGluAspGlyAspTrpPheThrValLeuSerGluValSerGlyArgGlyTr 80
        |||
Db      181  ATGCTCTCTGAGATGAGACTGAGTGGACGGGTGCTGTGAGTCTCAGGCGAGAGTAT 240
QY      81  AsnIleProSerValIleValAlaIysValSerHisGlyTrpLeuTyrlleGlyLeuSer 100
        |||
Db      241  AACATCCCACAGCTCCACAGTGCACAAATCTCCATGGGTGCTGATAGAGGCTTGAGC 300
QY      101  ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
        |||
Db      301  AGGGAAGAAAGCAGAGAGACTGCTGTGTTACTCTGGAACTCTGAGGGGCTTCCTCATC 360
QY      121  ArgGluSerGlnThrArgArgGlySerTyrlleSerLeuSerValArgLeuSerArgPro 140
        |||
Db      361  CGGAGAGCCAGACCCAGAGAGGCTTTACTCTGTCTGTCAGTCCGCTCAGCCGCTGCA 420
QY      141  SerTrpAspArgIleArgHisTyrlleArgIleHisCysLeuAspAsnGlyTrpLeuTyrlle 160
        |||
Db      421  TCCTGGAGCCGGATCAGACATCAGATCAGATCCATCCCTTGACATGCTGGCTGATCATC 480
QY      161  SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrlleSerGluLeuAla 180
        |||
Db      481  TCACCGGCGCCCTCACCTTCCCTCCTCAGGAGCCCTGGTGACCATTTACTGAGCTGGGG 540
QY      181  AspAspIleCysCysLeuLeuGluProCysValLeuGlnArgAlaGlyProLeuPro 200
        |||
Db      541  GATGACATCTCTGCTCTCACTAAGAGAGCCCTGTGCTGCGAGAGGCTGGCCGCTCCCT 600
QY      201  GlyLysAspIleProleuProValThrValGlnAqThrProleuAsnTrpIysGluLeu 220
        |||
Db      601  GGCAAGATATACCCCTACTGCTGAGTGCAGAGACACCACTCAACCTGGAAGAGCTG 660
QY      221  AASPserLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluGly 240
        |||
Db      661  GACAGCTCCCTCTCTTTCTGAGTGCACAGGGAGAGAGTCTCTTCTCAGTGAAGGT 720
QY      241  LeuArgGluSerLeuSerPheTyrlleSerLeuAsnAspGluAlaValSerLeuAsp 260
        |||
Db      721  CTCGGAGAGTCCCTCAGCTTCTACATCAGCTGAATGACGAGCTGTCTCTTGATAT 780
QY      261  Ala 261
        |||
Db      781  GCC 783

RESULT 4
AF326353      786 bp  mRNA  linear  PRI 08-NOV-2001
LOCUS        Homo sapiens Src-like adaptor protein-2 mRNA, complete cds.
DEFINITION   AF326353
ACCESSION    AF326353.1 GI:16797891
VERSION      AF326353.1
KEYWORDS
SOURCE
ORGANISM     Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 786)
AUTHORS      Holland,S.J., Liao,X.C., Mendenhall,M.K., Zhou,X., Pardo,J.,
              Chu,P., Spencer,C., Fu,A.C., Sheng,N., Yu,P., Pali,E., Nagin,A.,
              Shen,M., Yu,S., Chan,E., Wu,X., Li,C., Woiseleschlagier,M.,
              Ayers,G., Kolbinger,F., Bennett,M.K., Molineaux,S., Luo,Y.,
              Payan,D.G., Mancebo,H.S.Y. and Wu,J.
              Functional Cloning of Src-like Adaptor Protein-2 (SLAP-2), a Novel
              Inhibitor of Antigen Receptor Signaling
JOURNAL      J. Exp. Med. 194 (9), 1263-1276 (2001)

```

MEDLINE 21553259
 PUBMED 11696592
 REFERENCE 2 (bases 1 to 786)
 AUTHORS Holland, S.J., Mendenhall, M.K., Zhou, X., Spencer, C., Pardo, J., Fu, A.C., Sheng, N., Shen, M., Liao, C., Luo, Y., Payan, D.G., Mancho, H.S.Y. and Wu, J.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand Avenue, South San Francisco, CA 94080, USA
 FEATURES
 source
 1. 786
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 1. 786
 /note="SLAP2, SH2 and SH3 domain-containing adaptor"
 /codon_start=1
 /product="Src-like adaptor protein-2"
 /protein_id="FAL29204.1"
 /db_xref="GI:16797892"
 /translation="MGSLPSRRKSLPSPLSSVGGQGPVTWEAKSKATAVAGSPF AGPAPLSRLIGELTIVSDGDMWTVLSVSGREYNIPSVHAKVSHGMVYGLSRE KAEELLIPGNPGAFILRESQTRSGSYSLVRLSPASMDRIRHRIHCLDNGWLYI SPLTFPSLQALVDHYSRLADDICLLKEPCVLAGPLPKDIPPLVTVQRTPLNWK ELDLSLFSNAATGESLSLSEGLRESLSFYISLNDVSLDDA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.37e-101 Length: 786
 Score: 1351.00 Matches: 261
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 9
 US-10-043-649-2 (1-261) x AF326353 (1-786)

QY 181 ASPAPLIECYSCYSEULEUYSGLUPROCYVAILEUINALAGLYPROLEUPRO 200
 DB 541 GATGACATCTGCTGCTACTCAAGAGCCCTGTGTCTCGAGAGGCTGCCCTCCCT 600
 QY 201 GLYASPILEPROLEUPROVALTHVALGALNATH-PROLEUANTPLYSGLULEU 220
 DB 601 GGCAAGATATACCCCTACTGTGACTGTGCAGAGACACCACTCACTGAAAGAGCTG 660
 QY 221 ASPSERSEULEUPEHESERUALAATHGLYGLUJUSERLEUUSERGLUGLY 240
 DB 661 GACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGCTCTTCTCACTGAGGCT 720
 QY 241 LEUARGLUSERLEUSERPHERYILSERLEUASAPGLUALVALSERLEUASPAP 260
 DB 721 CTCGGAGAGCTCCTGACTTCAATCAAGAGCTGAAATGACAGAGCTGCTCTTGATGAT 780
 QY 261 ALA 261
 DB 781 GCC 783
 RESULT 5
 AK025645 2415 bp mRNA linear PRI 13-SEP-2003
 LOCUS Homo sapiens cDNA: FLJ21992 fls, clone HEP0554.
 DEFINITION AK025645
 ACCESSION AK025645.1 GI:10438227
 VERSION AK025645.1
 KEYWORDS oligo capping; fls (full) insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Kabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Obaishi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2415)
 Sugano, S., Suzuki, Y., Ota, T., Obaishi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
 Direct Submission
 Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fhdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
 NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
 FEATURES
 source
 1. 2415
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEP0554"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /clone_id="HEP"
 /note="cloning vector pME18SFL3"
 61. 846
 /note="unamed protein product"
 /codon_start=1
 /protein_id="BAB15201.1"
 /db_xref="GI:10438228"
 /translation="MGSLPSRRKSLPSPLSSVGGQGPVTWEAKSKATAVAGSPF AGPAPLSRLIGELTIVSDGDMWTVLSVSGREYNIPSVHAKVSHGMVYGLSRE KAEELLIPGNPGAFILRESQTRSGSYSLVRLSPASMDRIRHRIHCLDNGWLYI SPLTFPSLQALVDHYSRLADDICLLKEPCVLAGPLPKDIPPLVTVQRTPLNWK


```

/gene="SLA2"
/codon_start=1
/product="Src-like adaptor 2, isoform a"
/protein_id="AAH42041.1"
/db_xref="GI:27469843"
/db_xref="LOCUSID:84174"
/db_xref="MIM:606577"
/translation="MGSLSRRLKSLPSLSVSGOGPYTMEERSKATAVALGSEFP
AGGPALSLRLGEPRLTVSEDDGMVTLSPVSRVYNISPVYAKVSHCMVGLSRE
KAEELLLPKNRGGAFILRESQTRRSYSISLVSAPSPMDRIHRIRHICLDMGLYI
SPRLTPPSLQALVDVHSELADDICLLKEPCVQRAGPUPGKDIPLVTVQRTPLMK
ELDSILFSFSAATGESLISEGLRESLSFYISINDEAVSLDA"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 5,34e-101 Length: 2538
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

```

US-10-043-649-2 (1-261) x BC042041 (1-2538)

```

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
DB 363 ATGGAGAGTCTGCCAGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 422
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 423 CAAGGCCAGGAGCCTGTGACCATGAGACAGAGAGAGCAAGGCCACAGCCGTGGCCCTG 482
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 483 GGCAGTTTCCCGAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGAGAGCCATTGACC 542
QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyArgGluTyr 80
DB 543 ATGCTCTCTGAGAGTGAAGACTGGTGTGAGTGTCTGTAAGTCTCAGCGAGAGAGTAT 602
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrPleuTyrGluGlyLeuSer 100
DB 603 AACATCCCGACGCTCCACGTGGCCAAAGTCTCCCATGGGTGGCTGATGAGGAGCTGAGC 662
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 663 AGGAGAGAAAGCAGAGAACTGCTGTGTACCTGGAGAACTCTGGAGGGGCTTCTCTATC 722
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 723 CGGAGAGACCCAGACAGAGAGGCTTACTCTCTGTCAAGTCCGCTCAGCCGCTGCA 782
QY 141 SerTyrPheArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160
DB 783 TCCTGGGACCCGAGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGTGATCATC 842
QY 161 SerProAlaGluLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlyLeuAla 180
DB 843 TCACCGCCCTCACTTCCCTCACTCAAGGCCCTGTGGACCATTAATCTGAGAGTGGCG 902
QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
DB 903 GATGACATCTGCTGCTACTCAAGAGGCCCTGTGTCTGACAGAGGGCTGGCCGCTCCCT 962
QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
DB 963 GGCAGAGATATAACCTCACTGTGACTGTGACAGAGAACCACTCACTGAGAAAGCTG 1022
QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGlyGly 240
DB 1023 GACAGCTCTCTCTGTTTCTGAAAGCTGCACAGGAGAGAGTCTCTCTCTGATGAGGT 1082
QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260

```

```

DB 1083 CTCGGAGATCCCTCAGCTTCTACATCAGCTGAAATGAGAGGCTGTCTTTGATGAT 1142
QY 261 A1A 261
DB 1143 GCC 1145

```

```

RESULT 7
LOCUS AX452880
DEFINITION Sequence 1 from Patent WO0242457.
ACCESSION AX452880
VERSION AX452880.1 GI:21712520
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and
Kanner,S.B.

```

TITLE Cloning and expression of human slap-2: a novel sh2/sh3 domain-containing human slap homologue having immune cell-specific

```

JOURNAL Expression
Patent: WO 0242457-A 1 30-MAY-2002;
Bristol-Myers Squibb Co. (US)

```

FEATURES

```

source 1..2567
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 5,41e-101 Length: 2567
Score: 1351.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-10-043-649-2 (1-261) x AX452880 (1-2567)

```

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
DB 415 ATGGAGAGTCTGCCAGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTC 474
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
DB 475 CAAGGCCAGGAGCCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
DB 535 GGCAGTTTCCCGAGAGGTGGCCCGCCGAGCTGTGCTGAGACTCGGAGAGCCATTGACC 594
QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyArgGluTyr 80
DB 595 ATGCTCTCTGAGAGTGAAGACTGGTGGACGGTGTCTGTAAGTCTCAGGAGAGAGTAT 654
QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrPleuTyrGluGlyLeuSer 100
DB 655 AACATCCCGACGCTCCACGTGGCCAAAGTCTCCATGGGTGGCTGTATGAGGGCTGAGC 714
QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
DB 715 AGGAGAGAAAGCAGAGAACTGCTGTGTACTCTGGAACCTCGAGAGGGGCTTCTCTATC 774
QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
DB 775 CGGAGAGACCCAGACAGAGAGGCTTACTCTGTGTGATCGCGCTCAGCGCCCTGCA 834
QY 141 SerTyrPheArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160

```

Db 835 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGCTGATC 894
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 Db 895 TCACCGCGCTCCTACCTTCCCTCCTCCTCCTCCTCCTGAGCCCTGACCATTAATCTGAGCTGGCG 954
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
 Db 955 GATGACATCTGCTGCTTACTCAAGAGGCTTGTCTCTGACAGAGGCTGGCCGCTCCCT 1014
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
 Db 1015 GGCAAGAGATATACCCCTACCTGACTGTGACAGAGACCACTCACTGAAAGAGCTG 1074
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240
 Db 1075 GACACCTCCCTCTCTTTTCTGAAAGCTGCACAGAGGAGAGTCTTCTCTGAGTGGT 1134
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
 Db 1135 CTCGGGAGTCCCTCAGCTTCTACATCAAGCTGAATGAGAGGCTGTCTCTTGATGAT 1194
 QY 261 Ala 261
 Db 1195 GCC 1197

RESULT 8 AX443133 1183 bp DNA linear PAT 02-JUL-2002
 LOCUS AX443133
 DEFINITION Sequence 74 from Patent WO0216599.
 ACCESSION AX443133
 VERSION AX443133.1 GI:21690555
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Burgess, C.E., Conley, P.B., Grose, W.M., Hart, M., Kekuda, R.,
 Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,
 Topper, J.N. and Yang, R.B.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;
 Curagen Corporation (US) : COR THERAPEUTICS, INC. (US)

FEATURES
 source location/Qualifiers
 1..1183
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 4.68e-101 Length: 1183
 Score: 1347.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0
 Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 99.70% Indels: 0
 DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX443133 (1-1183)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
 Db 398 ATGGAGAGTTCGCCACGACAAAGAAATCTCTGCAAGCCCAAGCTTGAGTCTCTGTC 457
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 Db 458 CAAGGCAAGGAGACTGTGACCATGGAAGCAAGAAAGCAAGGCAAGCCGTGGCCCTG 517
 QY 41 GlySerPheProAlaGlyLysProAlaGluLeuSerLeuArgLeuGlyLysProLeuThr 60
 Db 518 GGCAAGTTCCCGGAGAGTGGCCCGGCGGAGCTGTCTGAGACTCGGAGGACCATGAGC 577

QY 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgGlyUyr 80
 Db 578 ATCGTCTCGAGAGATGGAGACTGCTGAGAGCGTGTCTGTGAAGTCTCAGGCAAGAGAT 637
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluLysSer 100
 Db 638 AACATCCCCAGCGTCACGTGGCAAGCTCCCATGAGGTGCTGTATGAGGGCTGAGC 697
 QY 101 ArgGlyLysAlaGluLeuLeuLeuLeuProGlyAsnProGlyValAlaPheLeuIle 120
 Db 698 AGGGAAGAACAGAGAACTGCTGTGTACTTGGAACCCCTGAGAGGGGCTTCTCTCANC 757
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 758 CGGAGAGCCAGAGCCAGAGAGGCTTACTCTGTGATGCTGCTGAGCCCTGCA 817
 QY 141 SerTrpAspAlaGlyLeuArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 Db 818 TCCTGGGACCGGATCAGACACTACAGATCCACTGCTTGACAAATGGCTGGCTGATC 877
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 Db 878 TCACCGCGCTCCTACCTTCCCTCCTCCTCCTCCTGAGCCCTGACCATTAATCTGAGCTGGCG 937
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
 Db 938 GATGACATCTGCTGCTACTCAAGAGCCCTGTCTCTGCAAGAGGCTGGCCGCTCCCT 997
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLeu 220
 Db 998 GGCAAGAGATATACCCCTACCTGACTGTGACAGAGACCACTCACTGAAAGAGCTG 1057
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGluGly 240
 Db 1058 GACACCTCCCTCTCTTTTCTGAAAGCTGCACAGAGGAGAGTCTTCTCTGAGTGGT 1117
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 260
 Db 1118 CTCGGGAGTCCCTCAGCTTCTACATCAAGCTGAATGAGAGGCTGTCTCTTGATGAT 1177
 QY 261 Ala 261
 Db 1178 GCC 1180

RESULT 9 AX443135/c 1183 bp DNA linear PAT 02-JUL-2002
 LOCUS AX443135
 DEFINITION Sequence 76 from Patent WO0216599.
 ACCESSION AX443135
 VERSION AX443135.1 GI:21690556
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1
 AUTHORS Burgess, C.E., Conley, P.B., Grose, W.M., Hart, M., Kekuda, R.,
 Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,
 Topper, J.N. and Yang, R.B.
 TITLE Proteins and nucleic acids encoding same
 JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;
 Curagen Corporation (US) : COR THERAPEUTICS, INC. (US)

FEATURES
 source location/Qualifiers
 1..1183
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 4.68e-101 Length: 1183
 Score: 1347.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62%
 Query Match: 99.70%
 DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX443135 (1-1183)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
 DB 786 ATGGAGAAAGTCTGCCAGCAGAAAGAAATCTTGCCAGGCCAAGCTTGAGTTCTCTGTC 727
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 DB 726 CAAAGCCAGAGAACCTGTGACATGAGAGCAGAGAGAACGACGACGCGGCGCTG 667
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 DB 666 GGCAGATTCCCGCAGATGCGCCGCGCGCTGCTGAGATCGGAGGACCATGACC 607
 QY 61 IleValSerGluAspGlyAspTyrPThrValLeuSerGluValSerGlyArgGluTyr 80
 DB 606 ATGCTCTGAGGATGAGACTGTGTGAGACGCTGCTGTGAATCTCAGGCAAGAGTAT 547
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
 DB 546 AACATCCCAAGCGTCCACGTCGGCAAGTCTCCATGCGTGTATGAGGCGCTGAGC 487
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 DB 486 AGGAGAAAGCAGAGAACTGCTGTTGTACCTGGGAACTTGAGAGGCGCTTCCTCATC 427
 QY 121 ArgGluSerGlnThrArgArgLysSerTyrSerLeuSerValArgLeuSerArgProAla 140
 DB 426 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTCTGCTGACCTCCGCTCAGCCCGCTGCA 367
 QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 DB 366 TCTCTGGGACCGGATCAGACACTCAGAGATCCACTGCTGACAAATGGCTGCTGTATCATC 307
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 DB 306 TCACCGGCGCTCACCTTCCCTCTCCTCAGGCGCTGTGACCACTTCTGAGCTGCGC 247
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
 DB 246 GATGACATCTGCTGCTCTACTCAGAGAGCCCTGTCTCTGCAAGGCGTGGCGCTCCT 187
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrPLeuGlyLeu 220
 DB 186 GGCAGAGATTAACCCCTAAGCTGTGCTGAGAGACACCACTCAAGTGAAGAGCTG 127
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGlyGly 240
 DB 126 GACAGCTCCCTCCGTTTCTGAGAGCTGCCACAGGAGGAGGCTCTTCTGAGTGAAGGT 67
 QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
 DB 66 CTCCGGAGAGTCCCTCAGCTTCTAATCAGCTGAATGACGAGGCGTCTCTTGATGATAT 7
 QY 261 Ala 261
 DB 6 GCC 4
 RESULT 10
 AX780857 2788 bp DNA linear PAT 14-JUL-2003
 LOCUS Sequence 3014 from Patent WO03039443.
 AX780857
 ACCESSION
 AX780857.1 GI:32697851
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

1 Haeflrich, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,
 Dugas, M., Eile, R., Broz, B. and Mergenthaler, S.

TITLE

Novel genetic markers for leukemias

JOURNAL

Patent: WO 03039443-A 3014 15-MAY-2003;
 Deutsches Krebsforschungszentrum (DE) ;
 Ludwig-Maximilian-Universitaet Muenchen
 PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)

FEATURES

source

1. 2788
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1 856-100 Length: 2788
 Score: 1345.00 Matches: 260
 Percent Similarity: 99.62% Conservative: 0
 Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 99.56% Indels: 0
 DB: 6 Gaps: 0

US-10-043-649-2 (1-261) x AX780857 (1-2788)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
 DB 387 ATGGAGAAAGTCTGCCAGCAGAAAGAAATCTTGCCAGGCCAAGCTTGAGTTCTCTGTC 446
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 DB 447 CAAAGCCAGAGAACCTGTGACATGAGAGCAGAGAGAACGACGCGGCGCTG 506
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 DB 507 GGCAGATTCCCGCAGATGCGCCGCGCGCTGCTGAGATCGGAGGACCATGACC 566
 QY 61 IleValSerGluAspGlyAspTyrPThrValLeuSerGluValSerGlyArgGluTyr 80
 DB 567 ATGCTCTGAGGATGAGACTGTGTGAGACGCTGCTGTGAATCTCAGGCAAGAGTAT 626
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
 DB 627 AACATCCCAAGCGTCCACGTCGGCAAGTCTCCATGCGTGTATGAGGCGCTGAGC 686
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 DB 687 AGGAGAAAGCAGAGAACTGCTGTTACTCTGCTGAGAACCTTGAGAGGCGCTTCCTCATC 746
 QY 121 ArgGluSerGlnThrArgArgLysSerTyrSerLeuSerValArgLeuSerArgProAla 140
 DB 747 CGGAGAGCCAGACAGAGAGGCTCTTACTCTGTGACATCCGCTTCAAGCGCGCTGCA 806
 QY 141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 DB 807 TCTGGAGACCGGATCAGACACTCAGAGATCCACTGCTTGAACATGGCTGATGACTC 866
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 DB 867 TCACCGGCGCTCACCTTCCCTCTCCTCAGGCGCTGTGACCACTTCTGAGAGTGAAGGT 926
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPro 200
 DB 927 GATGACATCTGCTGCTCTACTCAGAGAGCCCTGTGCTGCAAGGCGTGGCGCTCCT 986
 QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrPLeuGlyLeu 220
 DB 987 GGCAGAGATTAACCCCTAAGCTGTGCTGAGAGACACCACTCAAGTGAAGAGCTG 1046
 QY 221 AspSerSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuSerGlyGly 240
 DB 1047 GACAGCTCCCTCCGTTTCTGAGAGCTGCCACAGGAGGAGAGTCTTCTGAGTGAAGGT 1106

QY 241 LeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAsp 260
 Db 1107 CTCGGGAGTCCCTCAGCTTCTTACATCAGCTGAAATGAGAGGCTGTCTCTTGTGATGAT 1166
 QY 261 Ala 261
 Db 1167 GCC 1169

RESULT 11
 AX511155
 LOCUS AX511155 737 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 6 from Patent WO0242452.
 ACCESSION AX511155
 VERSION AX511155.1 GI:23392047
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Mcglade, J.C. and Loreto, M.P.
 TITLE Adapter gene
 JOURNAL Patent: WO 0242452-A 6 30-MAY-2002;
 The Hospital for Sick Children (CA)
 FEATURES
 source 1..737
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,71e-89 Length: 737
 Score: 1200.50 Matches: 241
 Percent Similarity: 92.40% Conservative: 2
 Best Local Similarity: 91.63% Mismatches: 1
 Query Match: 88.86% Indels: 19
 Gaps: 1

US-10-043-649-2 (1-261) x AX511155 (1-737)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
 Db 1 ATGGGAATCTGCCAGCAGAAAGAAATCTCTGCCAAGCCCAAGTTTATGTTCTCTCTC 60
 QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
 Db 61 CAAAGCCAGAGGAGCTGTGACCATGAGCAGAGCAAGCAAGCCAGCCAGCCCTG 120
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuAlaGluGlyGluProLeuThr 60
 Db 121 GGGAGTTTCCGGAGGTGGCCGGCCGAGCTGTGCTGAGACTCGGGAGCCATTGAC 180
 QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyArgGluTyr 80
 Db 181 ATCGTCTCTGAGATGAGAGATGAGTGTGCTGTCTGTGAAGTCTTACAGGAGAGAT 240
 QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
 Db 241 AACATCCCGAGCTCCAGCTGCGCAAGTCTCCCATGGGTGGTGTATGAGGGCTGAGC 300
 QY 101 ArgGluLysAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 Db 301 AGGAGAAAGCAGAGGAGTGTCTGTGTACCTGAGAACCTTGAGGGGCTTCTCATC 360
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 361 CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTGTGACGCGCCCTCAGCCGCTGCA 420
 QY 141 SerTyrAspArgIleAlaGlnIleTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
 Db 421 TCTCGGACCGGATCAGACTACAGATCCACTGCTTACATGCTGCTGTATCATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAla 180
 Db 481 TCACCGGCGCTACCTTCCCTCAGCTCCAGGCGCTGTGAGCCATTAC----- 528
 QY 181 AspAspIleCysCysLeuLeuLysGluProCysValLeuGln-ArgAlaGlyProLeuP 200
 Db 529 -----TGTAGGGCTGCGCCGCTCC 549

QY 200 OGlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrLysGluLe 220
 Db 550 TGGCAGAGATATACCCCTTACTGTGAGGAGGAGACCACTCATCTGAGAAAGAGCT 609

QY 220 uAspSerLeuLeuPheSerGluAlaAlaThrGlyGluGluSerLeuLeuSerGluG 240
 Db 610 GGACAGCTCCCTCCGTTTCTGAAAGCTCCACAGGAGAGAGTCTTCTTCAGTAGAG 669

QY 240 YLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAsp-GluAlaValSerLeuAsp 260
 Db 670 TCTCCGGAGTCCCTCAGCTTCTACATCAGCTGATAGAGGAGGCTGTCTTGGAGT 729

QY 260 SPAla 261
 Db 730 ATGCC 734

RESULT 12
 AF290986 737 bp mRNA linear PRI 21-JAN-2003
 LOCUS AF290986
 DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
 complete cds; alternatively spliced.
 ACCESSION AF290986
 VERSION AF290986.1 GI:17351922
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 737)
 Loreto, M.P. and Mcglade, C.J.
 TITLE Cloning and characterization of human Src-1-like adaptor protein 2
 JOURNAL Oncogene 22 (2), 266-273 (2003)
 MEDLINE 22415750
 PUBMED 12527895

REFERENCE
 2 (bases 1 to 737)
 Loreto, M.P. and Mcglade, C.J.
 TITLE Direct Submission
 JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
 FEATURES
 source 1..737
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="20"
 /tissue type="thymus"
 1..633
 /note="SLAP-2-v"
 /codon_start=1
 /product="Src-1-like adaptor protein-2 splice isoform"
 /protein_id="AA158198.1"
 /db_xref="GI:17351923"
 /translation="MGSLPSRRKSLPSLSVQGGPYTMEARSKATPAVALGSPF
 AGGPAELSLGELTIVSDGDMWTVLSVSGREYVIPSIVAKVSHGMYEGLSR
 KAEILLIPNGGAPLIRRSOTRRGSYSIVSRASRASMDRIRHRIHICIDNQMVI
 SPLTLPFSLDALVDHSSEMPARMQGTTPTCDCAEDDTQLERAGQLPVP"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,71e-89 Length: 737
 Score: 1200.50 Matches: 241
 Percent Similarity: 92.40% Conservative: 2
 Best Local Similarity: 91.63% Mismatches: 1
 Query Match: 88.86% Indels: 19

DB: 9 Gaps: 1

US-10-043-649-2 (1-261) x AF290986 (1-737)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20

DB 1 ATGGGAAAGTCTGCCAGAGAAAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC 60

QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerLysAlaThrAlaValAlaLeu 40

DB 61 CAAAGCCAGAGGACCTTGACCATGAGACAGAGAAAGCAAGGCCACAGCCGCTGAGCCCTG 120

QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60

DB 121 GCGAGTTTCCCGCAGGTGGCCCGCGGAGCTCTGCTGAGATTCGGGAGGACATTGACC 180

QY 61 IleValSerGlnAspGlyAspTrpTrpThrValLeuSerGlnValSerGlyArgGlnTyr 80

DB 181 ATCGTCTGAGATGAGACTGCTGAGACGCTGCTGTAAGTCTTCAGGACAGAGATAT 240

QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGlnGlyLeuSer 100

DB 241 AACATCCCCAGCGCTCCAGCTGGCCAAAGTCTCCCATGAGTGGCTGATAGAGGCTGAGC 300

QY 101 ArgGlnLysAlaGlnGlnLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120

DB 301 AGGAGAAAGACAGAGAACTGCTGTTGTTACCTGGGAAACCTGGAGGGGCTTCTCATC 360

QY 121 ArgGlnSerGlnThrArgArgLySerTyrSerLeuSerValArgLeuSerArgProAla 140

DB 361 CGGAGAGCCAGACCAAGAGAGGCTCTTACTCTCTGCAAGTCCGCTCACGCCGCTGCA 420

QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheTyrIle 160

DB 421 TCCGGAGACCGGATCAGACACTACAGATCCACTGCTTGACATGGCTGCTGATATC 480

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180

DB 481 TCACCGGCTCCTACCTTCCCTCATCTCCAGGCCCTGGTGGACCATTC- 528

QY 181 AspAspIleCysCysLeuLeuLysGlnProCysValLeuGln-ArgAlaGlyProLeuPr 200

DB 529 -----TCGAGGGCTGGCCGCTGCC 549

QY 200 CGLylLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlnLe 220

DB 550 TGGCAAGGATATACCTTACCTGACCTGCGAGGACACCACTCACTGAAAGACT 609

QY 220 uAspSerSerLeuLeuPheSerGlnAlaAlaThrGlyGlnGlnSerLeuLeuSerGlnI 240

DB 610 GGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGGAGAGTCTTCTCAGTGAAGG 669

QY 240 YLeuArgGlnSerLeuSerPheTyrIleSerLeuAsnAsp-GlnAlaValSerLeuAspA 260

DB 670 TCTCCGGAGAGCCCTCAGCTTCTACATCAGCTGATGAGGAGGAGGCTGCTCTTGAATG 729

QY 260 spAla 261

DB 730 ATGCC 734

RESULT 13

AX511151 777 bp DNA linear PAT 27-SEP-2002

LOCUS DEFINITION Sequence 2 from Patent WO0242452.

ACCESSION AX511151

VERSION AX511151.1 GI:23392045

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Mcglade, J.C. and Loreto, M.P.

TITLE Adapter gene

JOURNAL Patent: WO 0242452-A 2 30-MAY-2002;

FEATURES The Hospital for Sick Children (CH)

source location/Qualifiers

1..777

/organism="Mus musculus"

/mol_type="unassigned DNA"

/db_xref="taxon:10090"

ORIGIN

Alignment Scores: 1,83e-75 Length: 777

Pred. No.: 1032.00 Matches: 209

Score: 85.88% Conservative: 16

Percent Similarity: 79.77% Mismatches: 33

Best Local Similarity: 76.39% Indels: 3

Query Match: 6 Gaps: 3

DB: US-10-043-649-2 (1-261) x AX511151 (1-777)

QY 1 MetGlySerLeuProSerArgArgLySerLeuProSerProSerLeuSerSerVal 20

DB 1 ATGGGAAAGTCTGCCAGAGAGGAAACCC--TCCAGCCCCAGCCCCAGCTCTCTGAT 57

QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerLysAlaThrAlaValAlaLeu 40

DB 58 CCAAGCCAGAGAACCTGCTCATGACACAGAAAGACAAAGTCTACAGCTGAGCCCTG 117

QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60

DB 118 GCGAGTTTCCCGCAGGTGGCCCGCGGAGCTCTGCTGAGATTCGGGAGGACGCTGACC 177

QY 61 IleValSerGlnAspGlyAspTrpTrpThrValLeuSerGlnValSerGlyArgGlnTyr 80

DB 178 ATCATCTGAGATGAGATTGCTGAGACGCTCAGTCGAGAGTCTCAGGACAGAGATAC 237

QY 81 AsnIleProSerValHisValAlaLysValSerHisGlyTrpLeuTyrGlnGlyLeuSer 100

DB 238 CACATGCCCAAGTGTGATGAGCTTAAAGTCCGACAGGAGGCTGATCAGAGGCTGAGC 297

QY 101 ArgGlnLysAlaGlnGlnLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120

DB 298 CGGAGAAAGCCAGAGAACTACTCTGTTACTGGAACCCCGAGGGGCTTCTCATC 357

QY 121 ArgGlnSerGlnThrArgArgLySerTyrSerLeuSerValArgLeuSerArgProAla 140

DB 358 CGGAGAGCCAGACCAAGAGAGGCTGCTATTCCTGCTCCGCTCCAGACTCAGGCCCTGCA 417

QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPheTyrIle 160

DB 418 TCTTGGACCGGATCAGACACTACAGATCAGGCTCTTGACATATGCTGCTGATCATC 477

QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180

DB 478 TCACCTGCTCCTACCTTCCCTCATCTCCAGCCTTGGGAGGACATTCTGAGAGCTACCA 537

QY 181 AspAspIleCysCysLeuLeuLysGlnProCysValLeuGln-ArgAlaGlyProLeuPr 200

DB 538 GATGCACTGCTGCTGCTCCCTCAGGAGAGCGGTGTCTGCGAGAACTGGGCACTTACT 597

QY 201 GlyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlnLeu 220

DB 598 GGCAGAAATACCTCCACCTGTGACTGTGCCAACATCATCACTAATTTGAAAAAGCTG 657

QY 221 AspSerSerLeuLeuPheSerGlnAla--AlaThrGlyGlnGlnSerLeuSerGlnI 239

DB 658 GACCCGACCTCTCTGTTCTGGAAGCACCTGGAGTGGGAGGCACTCTGCTCAGTGAAG 717

QY 240 GlyLeuArgGlnSerLeuSerPheTyrIleSerLeuAsnAspGlnAlaValSerLeuAsp 259

DB 718 GCGCTCCAGAGTCCCTCAGTCTTCAATCAAGCTGCTGAGAGAC-----CCTTGAGAT 771

QY 260 spAla 261

Db 772 GATGCT 777

RESULT 14
AX511150 1348 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1 from Patent WO0242452.
ACCESSION AX511150
VERSION AX511150.1 GI:23392044
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Meglade, J.C. and Loreto, M.P.
TITLE Adapter gene
JOURNAL Patent: WO 0242452-A 1 30-MAY-2002;
The Hospital for Sick Children (CA)
FEATURES
source
1.1348
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN

Alignment Scores:
Pred. No.: 3.47e-75 Length: 1348
Score: 1032.00 Matches: 209
Percent Similarity: 85.88% Conservative: 16
Best Local Similarity: 79.77% Mismatches: 33
Query Match: 76.39% Indels: 4
DB: Gaps: 3

US-10-043-649-2 (1-261) x AX511150 (1-1348)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerSerVal 20
Db 282 ATGGAAATTGTTGTCCAGAGAGGAGAAACCC--TCCAGCCCCAGCCCTCTCTGCT 338
QY 21 GlnGlnGlnGlnProValThrMetGlnValGlnArgSerLysAlaThrAlaValAlaLeu 40
Db 339 CCAGACCCAGAAACCGGTGTCATGCAACCAAGAACCAAGTCACAGCCTGTGGCCCTG 398
QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlnProLeuThr 60
Db 399 GGCAGTTTCCAGAGGAGGACAGCCAGCTCATCTCGAGACTCGGGAGCCGCTGACC 458
QY 61 IleValSerGlnAspGlyAspTrpThrValLeuSerGlnValSerGlyArgGlnTyr 80
Db 459 ATCATCTCTGAGATGAGATTTGGTGGACACTCGAGAGTCTCAGGAGAGAGTAC 518
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTrpLeuTyrGlnGlyLeuSer 100
Db 519 CACATGCCCGAGTGTGTATGTGGTAAAGTGCACCGGTGGCTGTACGAGGCGCTGACC 578
QY 101 ArgGlnTyrAlaGlnGlnLeuLeuLeuProGlnAsnProGlnGlyAlaPheLeuIle 120
Db 579 CGGAGAAAGCCGAGGAGAACTACTCTGTACTCTGGGAAACCCGAGGGGCGCTTCTCATC 638
QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 639 CGGAGAGCCAGACCCAGAGAGGCTGCTATTCCTGTCCGCTCAGCCGCGCTGCA 698
QY 141 SerTyrAspArgIleArgHisTyrArgIleHisGlySerLeuAspAsnGlyTrpLeuTyrIle 160
Db 699 TCTTGGGACCGGATCAGACTACAGATACAGGCTTTGACATGCTGCTGTATCATC 758
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180
Db 759 TCACCTGCTCCTCCTCCCTCCTCCTCCTCCTCCTCCTGAGAGCTATCTCGAGCTGACA 818
QY 181 AspAspIleCysCysLeuLeuLysGlnProCysValLeuGlnArgAlaGlyProLeuPro 200

Db 819 GATGCACTGCTGCTCCCTCAGGAGCGCTGTGCTGTCTCCAGAGCTTGGCCACTACT 878

QY 201 GlyLysAspIleProLeuProValThrValGlnAlaGlyThrProLeuAsnTrpLysGlnLeu 220
Db 879 GCGAAAGATACACCTCCACCTCTGTGACTGTGACCAATGATCTCACTAAATGGAAAAAGCTG 938

QY 221 AspSerSerLeuLeuPheSerGlnAla---AlaThrGlnGlnGlnSerLeuLeuSerGln 239
Db 939 GACCGCAGCTCTCTGTTTCTGTGAGAGCACTGCGAGTGGGAGGACATCTGCTCACTGAG 998

QY 240 GlyLeuArgGlnSerLeuSerPheTyrIleSerLeuAsnAspGlnAlaValSerLeuAsp 259
Db 999 GGGCTCCAGAGAGTCCCTCAGTCTTACATCAGCTTGGCTGAGAGAC-----CCCTTGAT 1052

QY 260 AspAla 261
Db 1053 GATGCT 1058

RESULT 15
AF287467 1348 bp mRNA linear ROD 03-JUN-2002
LOCUS AF287467
DEFINITION Mus musculus Src-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF287467
VERSION AF287467.1 GI:17351918
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1348)
Loreto, M.P., Berry, D.M. and Meglade, C.J.
AUTHORS Functional cooperation between c-Cbl and Src-like adaptor protein 2 in the negative regulation of T-cell receptor signaling
TITLE Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
JOURNAL MEDLINE 22022020
PUBMED 12024036
REFERENCE 2 (bases 1 to 1348)
Loreto, M.P. and Meglade, C.J.
AUTHORS Direct Submission
TITLE Submitted (14-JUL-2000) Brain Tumour Research Centre, Hospital for Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada
JOURNAL
FEATURES
source
1.1348
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/note="SLAP-2"
/codon_start=1
/product="Src-like adaptor protein-2"
/protein_id="AAU38196.1"
/db_xref="GI:17351918"
/translation="MGSLSRGKTTSSPSSGPDDEPVSMOPEHRKVTAVAGSPFA
GEQARLSIRLGEPLTISDEGDWMTVQSEVSGNHYVYVAVKVAHGMVYEGISREK
AEELLPLGNPGAGFLIRSGTCYSVSRPSPMRIRHYRQRDNGLTYS
PRLTPSHLVHVEELADGICCPRLRPVLOKGLPEKDTPTPTVPTSSLNMK
LDRSLFLERPAEGASLUSEGURESLSTYSIAEDPLDA"
381..551
/note="Region: SH3 domain"
552..824
/note="Region: SH2 domain"
1301..1306

ORIGIN
polyA_signal

Alignment Scores:
Pred. No.: 3.47e-75 Length: 1348
Score: 1032.00 Matches: 209
Percent Similarity: 85.88% Conservative: 16
Best Local Similarity: 79.77% Mismatches: 33
Query Match: 76.39% Indels: 4
DB: Gaps: 3

US-10-043-649-2 (1-261) x AF287467 (1-1348)

```
QY      1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
      282 ATGGAGAAATTGTTCAGAGAGAGAAACC---TCCAGCCCAAGCCCGCTCTCTGCT 338
QY      21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerIysAlaThrAlaValAlaLeu 40
      339 CCAGACCAGAGAACCCGCTGTCATGCAACAGAAAGACACAGGTACAGCTGTGGCCCTG 398
QY      41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgGlnGlnProLeuThr 60
      399 GGGAGTTTCCAGCAGGTGAAACAGCCAGCTATCTCTGAGACTCGGGAGCCGCTGACC 458
QY      61 IleValSerGlnAspGlyAspTyrTyrThrValLeuSerGlnValSerGlyArgGlnTyr 80
      459 ATCATCTGAGAGATGAGATTGTGTGACAGTCCAGTCCGAAAGTCTCAGCAGAGATAC 518
QY      81 AsnIleProSerValHisValAlaIysValSerHisGlyTyrLeuTyrGlnGlyLeuSer 100
      519 CACATGCCCAAGTGTATGTGGCTAAAGTGCACAGGGTGGCTGTACAGAGGCCCTGAGC 578
QY      101 ArgGlnIysAlaGlnGlnLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
      579 CCGAGAAAGCCGAGAACTACTCTGTACTGTGGAAACCCGAGGGGCTTCTCATC 638
QY      121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
      639 CCGAGAGCCAGACACAGAGAGGCTGCTATTCCTGTCCTGCTCAGACTCAGCCGCTGCA 698
QY      141 SerTyrAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
      699 TCTTGGAGCCGATCAGACACTACAGATACAGGCTTGAACAATGGCTGCTGATCATC 758
QY      161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlnLeuAla 180
      759 TCACCTCGCCTCACCTTCCCTCCTCAGGCTTGGTGAAGCATTACTCTGAGCTAGCA 818
QY      181 AspAspIleCysCysLeuLeuGlnProCysValLeuGlnArgAlaGlyProLeuPro 200
      819 GATGGCATCTGCTGTCCCTCAGGAGCCGCTGTCTCTGCAAGACTTGGGCCACTACT 878
QY      201 GlyIysAspIleProLeuProValThrValGlnArgThrProLeuAsnTyrIysGlnLeu 220
      879 GGCAGAGATACACTCCACTGTGACTGTGCCAACATCATCATTAATTGGAAGAGCTG 938
QY      221 AspSerSerLeuLeuPheSerGlnAla---AlaThrGlnGlnGlnSerLeuLeuSerGln 239
      939 GACCGAGCCTCCTGTTCTGGAAGACACTGCAAGTGGGAGGCACTCTCTCAGTGAG 998
QY      240 GlnLeuArgGlnSerLeuSerPheTyrIleSerLeuAsnAspGlnAlaValSerLeuAsp 259
      999 GGGCTCCGAGAGTCCCTCAGTTCTTCACTCAGCCCTGCTGAGGAC-----CCCTGGAT 1052
QY      260 AspAla 261
      1053 GATGCT 1058
Db
```

Search completed: November 17, 2004, 02:15:10
Job time : 4555 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 17, 2004, 00:38:19 ; Search time 3371 Seconds

(without alignments)
2821.350 Million cell updates/sec

Title: US-10-043-649-2
Perfect score: 1351
Sequence: 1 MGSLSRRKSLPSPSLSSV.....RSLSPYISLNDPAVSLDDA 261

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 32822875 segs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-O=/cgn2.1/USPTO_spo01/US1003649/runat.16112004.060536.28728/app.query.fasta.1.455
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human0.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1003649@cgn2.1.1.3437 @runat.16112004.060536.28728 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1277	94.5	1002	5	BQ052308 AGENCOURT
2	1187	87.9	1069	5	BQ052468 AGENCOURT
3	1152.5	85.3	1020	5	BQ054281 AGENCOURT
4	1032	76.4	2637	3	AK088672 Mus muscu
5	1028	76.1	2974	3	AK030877 Mus muscu
6	937.5	69.4	926	3	AK020837 Mus muscu
7	882.5	65.3	660	6	BY742155 BY742155
8	871	64.5	1032	1	AL541041 AL541041
9	711	52.6	566	4	BG284179 BG284179

10	684.5	50.7	986	5	BQ054265	BQ054265 AGENCOURT
11	611.5	45.3	960	5	BU944126	BU944126 AGENCOURT
12	564.5	41.8	878	5	BQ053486	BQ053486 AGENCOURT
13	561	41.5	377	1	AA959151	AA959151 v251906.f
14	560.5	41.5	660	2	BB635615	BB635615 BB635615
15	555	41.1	597	1	AL844311	AL844311 AL844311
16	550	40.7	778	4	BG178487	BG178487 602328305
17	545	40.3	606	1	AL844309	AL844309 AL844309
18	543	40.2	614	1	AL844307	AL844307 AL844307
19	532	39.4	781	7	CK596391	CK596391 AGENCOURT
20	531	39.3	569	6	CB426333	CB426333 601508 MA
21	521.5	38.6	627	2	BB619854	BB619854 BB619854
22	520	38.5	762	5	BP162888	BP162888 BP162888
23	518	38.3	701	7	CK833360	CK833360 4057315 B
24	496	36.7	2810	3	AK036167	AK036167 Mus muscu
25	495.5	36.7	831	3	AY416279	AY416279 Mus muscu
26	495	36.6	1386	3	AK041565	AK041565 Mus muscu
27	493.5	36.5	831	9	AY416277	AY416277 Homo sapi
28	491	36.3	1762	3	CR617843	CR617843 full-length
29	482	35.7	1037	1	AL539427	AL539427 AL539427
30	481.5	35.6	655	7	CP913437	CP913437 A0648D04-
31	480.5	35.6	1997	3	AK037901	AK037901 Mus muscu
32	477	35.3	966	1	AL551370	AL551370 AL551370
33	474.5	35.1	701	5	EX436423	EX436423 BX436423
34	472	34.9	1015	5	EX436423	EX436423 BX436423
35	471	34.9	972	5	BQ707614	BQ707614 AGENCOURT
36	468	34.6	1054	1	AL549826	AL549826 AL549826
37	461	34.1	656	7	CN792083	CN792083 4126904 B
38	460	34.0	796	7	CO574805	CO574805 AGENCOURT
39	456.5	33.8	871	5	BQ436143	BQ436143 AGENCOURT
40	451.5	33.4	975	6	CB938697	CB938697 IPCGX13
41	448.5	33.2	934	7	CK411626	CK411626 AUF IPdHk
42	446	33.0	1063	5	BQ072745	BQ072745 AGENCOURT
43	445	32.9	723	5	BM950089	BM950089 UI-M-EHD
44	443.5	32.8	777	6	CB938723	CB938723 IPCGX13
45	443.5	32.8	882	5	BX415149	BX415149 BX415149

ALIGNMENTS

RESULT 1
LOCUS BQ052308
DEFINITION AGENCOURT 6868571 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933542
5', mRNA sequence.
ACCESSION BQ052308
VERSION BQ052308.1 GI:19811648
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1002)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L10M2118 row: d column: 23
High quality sequence stop: 670.
Location/Qualifiers
1. 1002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5933542"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 1,89e-116 Length: 1002
 Score: 1277.00 Matches: 257
 Percent Similarity: 97.74% Conservative: 2
 Best Local Similarity: 96.98% Mismatches: 2
 Query Match: 94.52% Indels: 4
 DB: 5 Gaps: 0

/tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: Blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

US-10-043-649-2 (1-261) x BQ052308 (1-1002)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
 Db 96 ATGGGAAAGTCTGCCAGACAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 155
 QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerlysalThrAlaValAlaLeu 40
 Db 156 CAAGGCCAGGAGACTGTGACCATGAGACAGAGAAAGCAAGGCCAGCGCTGGCCCTG 215
 QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60
 Db 216 GGCAGTTTCCCGAGAGTGGCCCGCCGAGCTGTGCTGACACTCGGAGACCATTAACC 275
 QY 61 IleValSerGlnAspGlyAspTyrThrValLeuSerGlnValSerGlyArgGlnTyr 80
 Db 276 ATGCTCTCTGAGATGAGACTGTGAGACGCTGTCTGAAGTCTCAGGCAGAGATAT 335
 QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrPleuTyrGlnGlyLeuSer 100
 Db 336 AACATCCCGACGCTCCACGTGGCCCAAAAGTCCCATGGGTGGCTGTATGAGGGCTGAGC 395
 QY 101 ArgGlnlysalGlnGlnLeuLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 Db 396 AGGAGAAAGCAGAGGAACTGCTGTGTACTGGAAACCTTGGAGGGGCTTCTCTATC 455
 QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 456 CGGAGAGCCAGACCCAGAGAGGCTCTACTCTGTCAATCCGCTCAGCCGCTGCA 515
 QY 141 SerTyrPaspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160
 Db 516 TCTTGAGACCCGATCAGACACTACAGATCTACTGCTTGCACAAAGGCTGGCTGATCATC 575
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGlyLeuAla 180
 Db 576 TCACCGCCCTCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 635
 QY 181 AspAspIleCysCysLeuLeuLeuGlnProCysValLeuGlnArgAlaGlyProLeuPro 200
 Db 636 GATGACATCTGCTGCTACTACTCAAGAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCT 695
 QY 201 GlnlysalPheProLeuProValThrValGlnArgThrProLeuAsnTyrlysalLeu 220
 Db 696 GGCAAGATATATCCCTTACTGTGACTGTGAGAGACACCACTCAACTGGAAGAAAGCTG 755
 QY 221 AspSerSerLeuLeuPheSerGlnAlaAlaThr-GlyGlnGlnSerLeuLeuSerGln-G 240
 Db 756 GACACCTCCTCTCTTTTCTGAACCTGCAAGGGAGGAGCTCTCTTCTTCTGAGAGG 815
 QY 240 lylenArg-GlnSerLeuSerPheTyrIleSerLeuAsnAspGln-AlaValSerLeuAs 259

Db 816 GTCTCCGGGAGATCCCTCAGCTTCTTACATCAGCCCTGATGACAGAGGCTGTCTTTGSA 875
 QY 259 paspAla 261
 Db 876 TGAATGCC 882

RESULT 2
 BQ052468 1069 bp mRNA linear EST 29-MAR-2002
 LOCUS AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
 DEFINITION 5' mRNA sequence.

ACCESSION BQ052468
 VERSION BQ052468.1 GI:19811808
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/MCI
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM2118 row: n column: 13
 High quality sequence stop: 681.
 Location/Qualifiers

FEATURES

1..1069
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5933772"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_106"
 /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1,93e-107 Length: 1069
 Score: 1187.00 Matches: 244
 Percent Similarity: 92.42% Conservative: 0
 Best Local Similarity: 92.42% Mismatches: 1
 Query Match: 87.86% Indels: 20
 DB: 5 Gaps: 1

US-10-043-649-2 (1-261) x BQ052468 (1-1069)

QY 1 MetGlySerLeuProSerArgArglySerLeuProSerProSerLeuSerSerVal 20
 Db 89 ATGGGAAAGTCTGCCAGACAGAAAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 148
 QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGlnArgSerlysalThrAlaValAlaLeu 40
 Db 149 CAAGGCCAGGAGCTGTGACCATGAGACGAGAGAGCAAGGCCAGCGCTGGCCCTG 208
 QY 41 GlySerPheProAlaGlyGlyProAlaGlnLeuSerLeuArgLeuGlyGlnProLeuThr 60
 Db 209 GGCAGTTTCCCGAGAGTGGCCCGCCGAGCTGTCTGAGACTCGGAGAGCATTAATGACC 268

```

QY 11LeValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db ATGCTCTCTGAGATGAGTGAAGTGTGTGAGCGTCTGTGAAAGTCTCAGGACAGAGATAT 328
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db AACATCCCAAGCGTCCACGTCGCGCAAGTCTCCCAAGGCGTGTATGAGGCGCTTGAGC 388
QY 101 ArgGluValAlaGluGluLeuLeuLeuLeuLeuProGlyAsnProGlyGlyValAphLeuLeu 120
Db AGGGAAGAAAGCAGAGAGAACTGCTGTGTATCACTGGAGACCTTGAGAGGCGCTTCCATC 448
QY 121 ArgGluSerGluThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db CCGGAGAGCCAGACCCAGAGAGGCTCTTACTCTGTCTGACGCTCGCTCAGCGCCCTGCA 508
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db TCCTGGAGCCGAGTACGACACTACAGAGATCCACTGCTTGAACATGCGCTGTACATC 568
QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeu-ValAspHisTyrSerGluLeuAl 180
Db 569 TCACCGCGCTCACCCTTCCCTCACCCTCAGCGCTGGGTGACCAATTACTCTGAGGG--- 625
QY 180 AspAspIleCysCysLeuLeuLysGluProCysValLeuGlnArgAlaGlyProLeuPr 200
Db 626 -----GCTGGCGCCCTCC 639
QY 200 GGIyLysAspIleProLeuProValThrValGlnArgThrProLeuAsnTrpLysGluLe 220
Db 640 TGGCAAGATATACCCCTACCTGACTGTGACAGGACACCACTCACTGAAAGAGCT 699
QY 220 LAspSerSerLeuLeuPheSerGluAlaAlaThr-GlyGluGluSerLeuSerGluG 240
Db 700 GGAACGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGGAGAGTCTTCTCACTGAGG 759
QY 240 LLeuArgGlu-SerLeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAsp 259
Db 760 GTCTCCGGGAAAGTCCCTCAGCTTCAATCAATGAGCCGTAATGAGAGGCTGTCTTTGAT 819
QY 260 AspAla 261
Db 820 GATGCC 825

```

RESULT 3
LOCUS B0054281 1020 bp mRNA linear EST 29-MAR-2002
DEFINITION AGNCOURT 6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
ACCESSION B0054281
VERSION B0054281.1 GI:19813621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LUCW2125 row: j column: 11
High quality sequence stop: 556.
Location/Qualifiers

```

source
1. 1020
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5936362"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 106"
/note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

Alignment Scores:
Pred. No.: 4,93e-104 Length: 1020
Score: 1152.50 Matches: 238
Percent Similarity: 93.51% Conservative: 7
Best Local Similarity: 90.84% Mismatches: 10
Query Match: 85.31% Indels: 7
DB: Gaps: 1

US-10-043-649-2 (1-261) x B0054281 (1-1020)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
Db 212 ATGGAAAGTCTGCGCCAGAGAAATCTCTGCGCAAGCCCAAGCTTGAATTCCTGTGC 271
QY 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerLysAlaThrAlaValAlaLeu 40
Db 272 CAGGCCAGGAGACCTGTGATCATGAGACAGAGAAACCAAGGCCACCGCTGGCCCTGC 331
QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db 332 GGCAGTTTCCCGGAGAGTGGGCCCGGCCAGCTGTGCTGAGACTGGGGAGCCATTGAC 391
QY 61 IleValSerGluAspGlyAspTrpThrValLeuSerGluValSerGlyArgGluTyr 80
Db 392 ATGCTCTCTGAGATGAGAGCTGAGTGGAGCGGTGCTGTGAAGTCTCAGGACAGAGAT 451
QY 81 AsnIleProSerValHisValAlaValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
Db 452 AACATCCCAAGCGTCCACGTCGCAAGTCTCCATGGGTGCTATGAGGCGCTTGAGC 511
QY 101 ArgGluValAlaGluGluLeuLeuLeuLeuLeuProGlyAsnProGlyGlyValAphLeuLeu 120
Db 512 AGGGAAGAAAGCAGAGAGAACTGCTGTGTATCACTGAGAACCTTGAGAGGCGCTTCTCATC 571
QY 121 ArgGluSerGluThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
Db 572 CGGAGAGCCAGACCCAGAGAGGCTTACTCTCTGTCAGTCCCTCAGCGCCCTGCA 631
QY 141 SerTrpAspArgIleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrLeuTyrIle 160
Db 632 TCCTGNAACCGGATCAACACTACAGATCCACTGCTTGAATGAGTGTGCTGTACATC 691
QY 161 SerProArgLeuThrPheProSerLeuGlnAla-LeuValAspHisTyrSerGluLeuAl 180
Db 692 TCACCGCGCTCACCCTTCCCTCAGCTCAAGCCCTGTGTGAGCAATTACTGTAGCTGCG 751
QY 180 AspAspIleCysCysLeu-LeuLysGluProCysValLeuGlnArgAlaGlyProLeu 200
Db 752 GGATGACATCTGCTGCTTACTGAGAGCCCTGTGTCTGCGCAAGGCGCTGCCGCTCC 811
QY 200 roGlyLysAspIle-ProLeuProValThrValGlnArgThrProLeuAsnTrpLysGlu 219
Db 812 CTGGCAAGATATACCCCTCACTGTGACTGTGCAAGAGACCACTCACTGGGAAAGAG 871
QY 220 LeuAspSerSerLeuLeu---PheSerGluAlaAlaThrGlyGluGluSerLeuSer 238

```

```

Db      872 CTGACGAGCTCCCTCTGTTTCTTGAGAGCTGCCCCAGGGAGAGAGTCTTTCATT 931
QY      239 GlnGly-LeuArgGlnSerLeuSerPheTyrIleSerLeu-AsnAspGluAla 255
      932 GAGGGGCTTCGCGGAGTCCCTCAGCTTCTACATCAACCTGTATGACGAGGCT 985

RESULT 4
AK086672
LOCUS   2637 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:EA30023D24 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.
ACCESSION AK086672
VERSION   1 GI:26353729
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS  Carninci, P. and Hayashizaki, Y.
TITLE    High-efficiency full-length cDNA cloning
JOURNAL  Meth. Enzymol. 303, 19-44 (1999)
MEDLINE  99279253
PUBMED   10349636

REFERENCE
AUTHORS  Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
          Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE    Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL  Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE  20499374
PUBMED   11042159

REFERENCE
AUTHORS  Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
          Komno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
          Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
          Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
          Fujisake, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matshiki, M.,
          Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J.,
          Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE    RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multichannel sequencer
JOURNAL  Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE  20530913
PUBMED   11076861

REFERENCE
AUTHORS  The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
TITLE    Functional annotation of a full-length mouse cDNA collection
JOURNAL  Nature 409, 685-690 (2001)
MEDLINE  11701259
PUBMED   11701259

REFERENCE
AUTHORS  The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE    Analysis of the mouse transcriptome based on functional annotation
          of 60,770 full-length cDNAs
JOURNAL  Nature 420, 563-573 (2002)
MEDLINE  12001259
PUBMED   12001259

JOURNAL
AUTHORS  Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
          Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
          Hiyama, S., Furuno, M., Hiramoto, K., Hiraoka, T., Hirozane, T.,
          Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kankawa, T.,
          Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M.,
          Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
          Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohseto, N.,
          Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
          Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shitaka, T.,
          Soabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
          Tanaka, Y., Tanaka, T., Tomaru, A., Toyo, T., Yasunishi, A.,
          Muramatsu, M. and Hayashizaki, Y.
TITLE    Direct Submission
JOURNAL  Submitted (16-APR-2002) Yoshinide Hayashizaki, The Institute of

```

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>
 URL: <http://fantom.gsc.riken.jp/>
 Location/Qualifiers

FEATURES

source

1.2637

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="PANTOM,DB:EA30023D24"

/db_xref="taxon:10090"

/clone="EA30023D24"

/cell_type="thymic cells"

/tissue_type="thymus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="2 days neonate"

358..1137

/note="unlabeled protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTF) AL38196, evidence: FASTY, 100%ID, 100%length, match=777)

putative"

/codon_start=1

/protein_id="BA040495.1"

/db_xref="GI:26353720"

/translation="MWSLSRSGKTSFSPSSGPDGPVSMQERHVTVAVALGPPA
 GEQARSLRGLPGLTIISEDDMTVQVSESGEYHMPSVYAKVAHGMVLEGLSRK
 AEELILGNPGAFILRESQTRGYSLSVRSLRSPMSRIRYRIQRLDNGWLVS
 PLTFPSHILVHVESLADGICCPLEPCVLOKLGPLPGKDPPTVPTVPSINWKK
 LNSILFLKAPASGKSLISGLRESLSYISLAEDPLDA"

CDS

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score: 1032.00	2637	209	16	33	4
Percent Similarity: 85.88%					
Best Local Similarity: 79.77%					
Query Match: 76.39%					

US-10-043-649-2 (1-261) x AK086672 (1-2637)

```

QY      1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
Db      358 ATGGAGATTGTGTCACGACGAGGAAACCC--TCCACCCCGAGCCGCTCTCTGCT 414
QY      21 GlnGlyGlnGlyProValThMetGluAlaGluArgSerIysAlaThrAlaValAlaLeu 40
Db      415 CCAGACCGGAAACCGGTCTCATGACAAACAAACAGCGTACAGCTGTGGCCCTG 474
QY      41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
Db      475 GGCAGTTTCCAGACGAGTGGAACAGGCCAGACTATCTCGAGCTCGGGAGCGCTGACC 534
QY      61 IleValSerGluAspGlyAspTyrPheThrValLeuSerGluValSerGlyArgGluTyr 80
Db      535 ATCATCTCTGAGGATGAGGATGTGTGACAGTCCAGTGGAGAGTCTCAGCGAGAGTAC 594
QY      81 AsnIleProSerValHisValAlaValSerHisGlyTyrPheTyrGlnGlyLeuSer 100
      .....

```


BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 9825.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?cs=CS0DE005AF120P1ec=9825.r.

FEATURES
 source
 Location/Qualifiers
 1..1032
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE005YK23"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with NotI and EcoRV sites of the
 pCMVSPORT 6 vector.
 the NotI and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 5,59e-76 Length: 1032
 Score: 871.00 Matches: 169
 Percent Similarity: 96.02% Conservative: 0
 Best Local Similarity: 96.02% Mismatches: 7
 Query Match: 64.47% Indels: 0
 DB: 1 Gaps: 0
 US-10-043-649-2 (1-261) x AL541041 (1-1032)

QY 1 MetGlySerLeuProSerArgArgLysSerLeuProSerProSerLeuSerSerVal 20
 Db 395 ATGGGAAGTCTGCCAGAGAAAGAAATCTCTGCCAAGCCCAAGTTAGTTCTCTCTTC 454
 QY 21 GlnGlyGlnGlyProValThrMetGlnAgluArgSerLysAlaThrAlaValAlaLeu 40
 Db 455 CAAAGCCAGGAGCCGTGACCATGAGAGAGAAAGCAAGGCCACGCTGGCCCTG 514
 QY 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuAgluGlyGluProLeuThr 60
 Db 515 GGCAGTTTCCCGGAGGTGGCCCGCCGCACTGTGCTGAGACTCGGGAGCCCATTTGACC 574
 QY 61 IleValSerGluAspGlyAspTyrTyrThrValLeuSerGluValSerGlyAlaArgGluTyr 80
 Db 575 ATCGTCTCTGAGATGAGACTGGTGGACGCTGCTGTAAGTCTCAGGAGAGAGATAT 634
 QY 81 AsnIleProSerValHisValAlaIleValSerHisGlyTyrPleuTyrGluGlyLeuSer 100
 Db 635 AACATCCCAAGCCGTCACGTCGCAAGTCTCCCATGAGTGGTGTATGAGAGGCTTGAAGC 694
 QY 101 ArgGlnLysAlaGlnGluLeuLeuLeuLeuProGlyLysAsnProGlyGlyAlaPheLeuIle 120
 Db 695 AGGAGAAAGAGAGAGACTGCTGTTTGTACCTGGAAACCTGGAGGCGCTTCTTCATC 754
 QY 121 ArgGlnSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 755 CGGAGAGCCAGAACAGAGAGGCTTACTCTCTGTCAAGTCCGCCCTCAGCGCCCTGCA 814
 QY 141 SerTrpAspArgIleAargHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIle 160
 Db 815 TCTTGGAGCCGAGACACTMAAGATCCACTGCTTGAACAATGCTGGCTGTATCATC 874
 QY 161 SerProArgLeuThrPheProSerLeuGlnAlaLeuValAspHisTyr 176
 Db 875 TCAAGGCGCTTAACTTCCCTCACTCCAGGCTGAGTGGACMATTTAC 922
 RESULT 9
 BG284179

LOCUS
 DEFINITION
 EC284179 566 bp mRNA linear EST 21-FEB-2001
 602408226F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4520382 5',
 mRNA sequence.
 ACCESSION
 BG284179
 VERSION
 BG284179.1 GI:13034866
 KEYWORDS
 EST.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 1 (bases 1 to 566)
 NIH-MGC http://mgi.mci.nih.gov/.
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Unpublished (1999)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: DCTP/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHM10418 row: c column: 07
 High quality sequence start: 2
 High quality sequence stop: 566.
 Location/Qualifiers
 1..566
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4520382"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="RDH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally, oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 2.02e-60 Length: 566
 Score: 711.00 Matches: 136
 Percent Similarity: 99.27% Conservative: 0
 Best Local Similarity: 99.27% Mismatches: 1
 Query Match: 52.63% Indels: 0
 DB: 4 Gaps: 0
 US-10-043-649-2 (1-261) x BG284179 (1-566)

QY 125 ThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAlaSerTyrAspArg 144
 Db 4 ACGGTCGCGGCTTAACTCTCTGTCACTCCGCTCAGCGCCCTGCATCTCGGAGACGG 63
 QY 145 IleArgHisTyrArgIleHisCysLeuAspAsnGlyTyrPleuTyrIleSerProArgLeu 164
 Db 64 ATCAGACACTACAGAGATCCACTGCTTGAACAATGGCTGTGATCACTCAGCGCTTC 123
 QY 165 ThrPheProSerLeuGlnAlaLeuValAspHisTyrSerGluLeuAlaAspAspIleCys 184
 Db 124 ACCTCCCTCACTCCAGCGCCCTGTGTGACCATTACTCTGAGCTGGCGGAGATATCGC 183
 QY 185 CysLeuLeuArgGluProCysValLeuGlnArgAlaGlyProLeuProGlyLysAspIle 204
 Db 184 TGCTTACTCAAGAGCCCTGTGTGTCAGAGAGGCTGGCCGCTCCCTGGCAAGATATA 243
 QY 205 ProLeuProValThrValGlnArgThrProLeuAsnTyrPleuGluLeuAspSerLeu 224
 Db 244 CCCCTACTGTGACTGTGAGAGAGACACCACTAACTGGAAGAGCTGGACAGCTCCCTC 303
 QY 225 LeuPheSerGlnAlaIleThrGlyGluGluSerLeuLeuSerGluGlyLeuArgGluSer 244

Db 304 CTTGTTTTCAGAGCTGCCAGAGGAGGCTCTTTCTAGTGAAGGCTCCGGAGTCC 363

Qy 245 LeuSerPheTyrIleSerLeuAsnAspGluAlaValSerLeuAspAspAla 261

Db 364 CTCAGCTTCTACATCAGCTGATGATGACGAGGCTGCTCTTGTGATGATGCC 414

RESULT 10

LOCUS BQ054265 986 bp mRNA linear EST 29-MAR-2002

DEFINITION AGENCOURT 6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339

ACCESSION BQ054265

VERSION BQ054265.1 GI:19813605

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 986)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLCM215 row: i column: 12
High quality sequence stop: 515.

FEATURES

source

1..986

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5936339"

/tissue_type="natural killer cells, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC_106"

/note="Organ: blood; Vector: pOT81; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2e-57 Length: 986

Score: 684.50 Matches: 164

Percent Similarity: 73.57% Conservative: 3

Best Local Similarity: 72.25% Mismatches: 16

Query Match: 50.67% Indels: 44

DB: 5 Gaps: 3

US-10-043-649-2 (1-261) x BQ054265 (1-986)

Qy 1 MetGlySerLeuProSerArgArgIleGlySerLeuProSerProSerLeuSerSerVal 20

Db 279 ATGGAAATCTGCGCCAGCAAGAAATCTCTGCCAAGCCCAAGCTTGATTCCTCTGTC 338

Qy 21 GlnGlyGlnGlyProValThrMetGluAlaGluArgSerIleValaThrAlaValAlaLeu 40

Db 339 CAAGGCCCGGAGCCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398

Qy 41 GlySerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60

Db 399 GGCAGTTTCCCGCAGGTGGCCCGGCCGAGCTGTGCTGAGACTCGGGAGGCAATTGACC 458

Qy 61 IleValSerGluAspGlyAspTrpTrpThrValLeuSerGluValSerGlyArgIleTyr 80

Db 459 ATCGTCTCTGAGAGTGAAGACTGCTGAGACCGTGTCTCTGAACTCTCAGGACAGAGTAT 518

Qy 81 AsnIleProSerValHisValAlaIleValSerHisGlyTyrLeuTyrGluGlyLeuSer 100

Db 519 AACATCCCAAGCGTCCACGCTGGCCAAAGTCTCCCATGGGTGCTGTATAGAGCCCTGAGC 578

Qy 101 ArgGluValAlaGluGluLeuLeuLeuLeuProGlyAsnProGlyValAla-PheLeuIle 120

Db 579 AGGAGAAAGCAAGAGAACTGCTGTGTACTGCTGAGAACCTGGAGAGGCGCTTCTCTCAT 638

Qy 120 eArgGlu-SerGlnThrArgArgIleSerTyrSerLeuSerVal-ArgLeuSerArg-Pr 139

Db 639 CCGGAGAAAGCCAGACCAAGAGAGAGCTTACTCTGTGACGTCCGCTCAGCCGCCCC 698

Qy 139 oAlaSerTrp-AspArg-IleArgHisTyrArg-IleHisCysLeuAspAsnGlyTyrPhe 158

Db 699 TGCATCTCTGGGACCGGGATCAGACCTTACAGGGATTCCCTTGTGAACCATTTGGCT 758

Qy 158 v-----TyrIleSerProArgLeuThrPheProSerLeuGlnAlaLeuValas 174

Db 759 TGGCTTGTAATATTTTAAACCGGGCTTACCTTTTCC----- 798

Qy 174 pHISrYsSerGluLeuAlaAspAspIleCysCysLeuLeuGluProCysValLeuGl 194

Db 799 -----CTTAA 803

Qy 194 nArgAlaGlyProLeuProGlyLysAspIleProLeuProValThrValGlnArgTrp 214

Db 804 ATTTCAAGGGGCCCTTGGGGGAACCATTT-----ACTCC 839

Qy 214 cLeuAsn 216

Db 840 TTTAAAC 846

RESULT 11

LOCUS BU944126 960 bp mRNA linear EST 18-OCT-2002

DEFINITION AGENCOURT_10545003 NIH_MGC_107 Homo sapiens cDNA clone

ACCESSION BU944126

VERSION BU944126

KEYWORDS IMAGE:6728350 5', mRNA sequence.

SOURCE BU944126.1 GI:24132945

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 960)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLCM3049 row: m column: 21
High quality sequence stop: 628.

FEATURES

source

1..960

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6728350"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

10 (11), 1757-1771 (2000)
 Komoto, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Suganara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, U., Shibata, K., and Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

Location/Qualifiers
 1..660
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="A430076N07"
 /tissue_type="thymus"
 /dev_stage="0 day neonate"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 0 day neonate thymus"
 /note="Site 1: SalI; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGGAGGAGAGATCCAGAGCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGGAGATTCGAGTAAATTAATTCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I."

ORIGIN

Alignment Scores:

Pred. No.: 2.51e-45 Length: 660
 Score: 560.50 Matches: 115
 Percent Similarity: 87.23% Conservative: 8
 Best Local Similarity: 81.56% Mismatches: 17
 Query Match: 41.49% Indels: 1
 DB: 2 Gaps: 1

US-10-043-649-2 (1-261) x BB635615 (1-660)

QY 1 MetGlySerLeuProSerArgArgGlySerLeuProSerProSerLeuSerSerVal 20
 Db 241 ATGGGAGTTGTTCACAGGAGGAAAAAC---TCCAGCCCCAGCCCCAGCTCTCGGT 297
 QY 21 GlnGlyGlnGlyProValThrMetGlnAlaGluArgSerIysAlaThrAlaValAlaLeu 40
 Db 298 CCAGACCAAGAACCCGTCATCATGCAACCAAGAACCAAGTCAGAGCTGGCCCTG 357
 QY 41 G1SerPheProAlaGlyGlyProAlaGluLeuSerLeuArgLeuGlyGluProLeuThr 60
 Db 358 GGCAGTTTCCAGAGGTGAACAGCCAGCACTATCTCTGAGACTCGGGAGCCCTGACC 417
 QY 61 IleValSerGluAspGlyAspTyrThrValLeuSerGluValSerGlyValArgGlyTyr 80
 Db 418 ATCATCTCTAGAGATGAGATGTGTGACAGCTCCAGAGGAGCTCAGGCAAGAGTAC 477
 QY 81 AsnIleProSerValHisValAlaIysValSerHisGlyTyrLeuTyrGluGlyLeuSer 100
 Db 478 CACATGCCCGAGTGTATGTGTAAAGTCGCCCAAGCGGTGGCTGTACAGAGGCGCTGAGC 537

QY 101 ArgGluTyrAlaGluGluLeuLeuLeuProGlyAsnProGlyGlyAlaPheLeuIle 120
 Db 538 CCGAGAAAGCCGAGGAATCTACTCTTACTCTGACACCCCGAGGAGGCTTCTCATC 597
 QY 121 ArgGluSerGlnThrArgArgGlySerTyrSerLeuSerValArgLeuSerArgProAla 140
 Db 598 CCGAGAGCCAGACCGAGAGAGGCTGTATTCCTCTCCGTTGATCAGCCGCTGCA 657
 QY 141 Ser 141
 Db 658 TCT 660

RESULT 15
 ALB44311 597 bp mRNA linear EST 30-JUL-2002
 LOCUS ALB44311 pool_YT_11b_v_SPD Homo sapiens cDNA, mRNA sequence.
 DEFINITION ALB44311
 ACCESSION ALB44311.1 GI:22019093
 VERSION ALB44311.1
 KEYWORDS EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 597)
 Ashcroft, K., Bethel, G., Bye, J. M., Howell, G. R., Huckle, E. J. and Sheridan, E.
 TITLE Homo sapiens EST sequence
 JOURNAL Unpublished (2002)
 COMMENT Contact: The Sanger Centre
 The Sanger Centre
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: humquerry@sanger.ac.uk
 Sanger Centre name: sc0d10818.154136A
 Homo sapiens EST sequence. This sequence was generated as part of The Wellcome Trust Sanger Institute program to identify and annotate genes in the human genome. Incomplete or unconfirmed genes are experimentally analysed using a variety of cDNA library resources. This sequence was obtained from a PCR product generated from a pool of up to 100,000 cDNA clones derived from pool_YT_11b_v_SPD cDNA library. Further information can be found at <http://www.sanger.ac.uk/Teams/Team69/>.

FEATURES

source

1..597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="20"
 /clone_lib="pool_YT_11b_v_SPD"

ORIGIN

Alignment Scores:

Pred. No.: 7.64e-45 Length: 597
 Score: 555.00 Matches: 107
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 41.08% Indels: 0
 DB: 1 Gaps: 0

US-10-043-649-2 (1-261) x ALB44311 (1-597)

QY 155 AsnGlyTyrPheTyrIleSerProArgLeuThrPheProSerLeuGlnAlaIleValAsp 174
 Db 13 AATGGCTGGCTGTACATCTCAACACGCTCACTTCCCTCACTCCAGGCCCTGTGTGAC 72
 QY 175 HisTyrSerGluLeuAlaAspPheIleCysCysLeuLeuIysGluProCysValIleGln 194
 Db 73 CATTACTCTGAGCTGGCGGATGACATCTGCTCACTCAAGAGAGCCCTGTGCTCAG 132
 QY 195 ArgAlaGlyProLeuProGlyIysAspIleProLeuProValThrValGlnAlaGlyThrPro 214
 Db 133 AGGGCTGGCCGCTCCTCGCAAGAGATATACCCCTACTGTGACTGTGACAGAGACACCA 192

```
QY      215 LeuAentRPLySGluLeuAspSerSerLeuLeuPheSerGluAlaAlaThrGlyGlu 234
      |||
Db      193 CTCAACTGGAAAGAGCTGACAGCTCCCTCCTGTTTCTGAAGCTGCCACAGGGAGAG 252
      |||
QY      235 SerLeuLeuSerGluGlyLeuArgGluSerLeuSerPheTyrIleSerLeuAsnAspGlu 254
      |||
Db      253 TCTCTTCTCAGTGAAGGCTCTCCGGAGTCCCTCAGCTTACATCAGCCTGAATGACGAG 312
      |||
QY      255 AlaValSerLeuAspAspAla 261
      |||
Db      313 GCTGCTCTTTGGATGATGCC 333
      |||
```

Search completed: November 17, 2004, 03:11:31
Job time : 3383 secs

QY 540 GGATGACATCTGCTGCTTA-CTCAAGAGCCCTGTGCTGAGAGGCGCCGCTCC 598
 Db 752 GGATGACATCTGCTGCTTANCTCAGGAGCCCTGTCTCTGCAAGAGGCTGCCGCTCC 811
 QY 599 CTGGCAGAGATATA-CCCTACTGTGACTGTGAGAGACACCACTCACTGAAAGAG 657
 Db 812 CTGGCAAAATATACCCCTACTGTGACTGTGCAAAAGACACCACTCACTGAAAGAG 871
 QY 658 CTGACAGCTCCCTCTG-TTTTCTGAGCTGCCCA-GGGAGAGCTCTTCTGACT 714
 Db 872 CTGGACAGTCCCTCTGCTGTTTCTTGAAGCTGCCCAAGGAGAGCTCTTCTGACT 931
 QY 715 GAGGG-TCCTCCGAGTCCCTCAGCTTACATCAGCCTG-AATGAGAGGCTGCTCT 771
 Db 932 GAGGGGCTCTCCGAGAGTCCCTCAGCTTACATCAGCCTGTAATGACAGGCTGCTCT 991
 QY 772 TTGG 775
 Db 992 TTGG 995

RESULT 3
 BQ052468 1069 bp mRNA linear EST 29-MAR-2002
 LOCUS DEFINITION AGENCOURT 6868422 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5933772
 ACCESSION BQ052468
 VERSION BQ052468.1 GI:19811808
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgrabbs-remail.nih.gov
 Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LNCM2118 row: n column: 13
 High quality sequence stop: 681.
 Location/Qualifiers

FEATURES

1..1069
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5933772"
 /tissue_type="natural killer cells, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_106"
 /note="Organ: blood; Vector: pOT7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."

ORIGIN

Query Match 82.1%; Score 645; DB 5; Length 1069;
 Best Local Similarity 93.4%; Pred. No.2.4e-159;
 Matches 737; Conservative 0; Mismatches 0; Indels 52; Gaps 4;
 QY 1 ATGGGAAGTGTGCCGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 60
 |||||

Db 89 ATGGGAAGTGTGCCGACAGAGAAATCTCTGCCAAGCCCAAGCTTGAATCTCTGTC 148
 QY 61 CAAGCCAGAGGACCTGTGACATGGAAGAGAGAGAAAGCCACAGCCCTGCGCTTG 120
 Db 149 CAAGCCAGAGGACCTGTGACATGGAAGAGAGAGAAAGCCACAGCCCTGCGCTTG 208
 QY 121 GGCAGTTTCCCGGAGGTGGCCCGGCGGAGCTGTGCTAGACTGCGGAGAGCATTAAGCC 180
 Db 209 GGCAGTTTCCCGGAGGTGGCCCGGCGGAGCTGTGCTAGACTGCGGAGAGCATTAAGCC 268
 QY 181 ATTCCTCTGAGAGATGAGACTGTGTGACAGGTGCTGTCTGAAATCTCAAGCAGAGAT 240
 Db 269 ATTCCTCTGAGAGATGAGACTGTGTGACAGGTGCTGTCTGAAATCTCAAGCAGAGAT 328
 QY 241 AACATCCCGAGGTCACAGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 300
 Db 329 AACATCCCGAGGTCACAGTGGCCCAAGTCTCCCATGGGTGGCTGTATGAGGCTGAGC 388
 QY 301 AGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAAACCTGAGAGGCGCTTCTCATC 360
 Db 389 AGGAGAAAGCAGAGAACTGCTGTGTACCTGGGAAACCTGAGAGGCGCTTCTCATC 448
 QY 361 CGGAGAGCCGACACCCAGAGAGGCTTTATCTCTGTGACAGTCCGCTCAAGCCCTGCA 420
 Db 449 CGGAGAGCCGACACCCAGAGAGGCTTTATCTCTGTGACAGTCCGCTCAAGCCCTGCA 508
 QY 421 TCCTGGAGCCGATCAGACATCAGAGATCGACTGCTTGAACATAGCTGGCTGTACATC 480
 Db 509 TCCTGGAGCCGATCAGACATCAGAGATCGACTGCTTGAACATAGCTGGCTGTACATC 568
 QY 481 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTT-GGTGAGCATTAATCTGAGAGTGGC 539
 Db 569 TCACCGGCGCTCACCTTCCCTCACTCCAGGCGCTTGGTGAACATTAATCTGAG- 623
 QY 540 GGATGACATCTGCTGCTTAACAGAGCCCTGTGCTGAGAGGCTGCGCTCCG 599
 Db 624 -----GGGCTGGCCCGCTCCG 639
 QY 600 TGGCAAGATATACCCCTACCTGACCTGACGTCGACAGAGACACCACTCACTGAAAGAGCT 659
 Db 640 TGGCAAGATATACCCCTACCTGACCTGACGTCGACAGAGACACCACTCACTGAAAGAGCT 699
 QY 660 GAGCAGCTCCCTCTGTTTCTGAGCTGCCACA-GGGAGAGAGTCTTCTGAGTGGAG 718
 Db 700 GAGCAGCTCCCTCTGTTTCTGAGCTGCCACA-GGGAGAGAGTCTTCTGAGTGGAG 759
 QY 719 GTCTCCGGG-AGTCCCTACGCTTCAATACAGCTGAATACAGAGCTGTCTTTGAT 777
 Db 760 GTCTCCGGGAGTCCCTACGCTTCAATACAGCTGAATACAGAGCTGTCTTTGAT 819
 QY 778 GATGCTAG 786
 Db 820 GATGCTAG 828

RESULT 4
 AK088672 2637 bp mRNA linear HTC 03-APR-2004
 LOCUS DEFINITION Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
 full-length enriched library, clone:E430023D24 product:MODULATOR OF
 ANTIGEN RECEPTOR SIGNALING MARS, full insert sequence.
 ACCESSION AK088672
 VERSION AK088672.1 GI:26353729
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253

PUBMED
 REFERENCE
 AUTHORS
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1650 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, U.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multipipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 6 (bases 1 to 2637)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Harai, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kadawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ono, N., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, P., Shibata, K., Shikigawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Substitution
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 220-0045, Japan (E-mail: genome-resgsc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 COMMENT
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
 Trust/MRC building Addenbrookes Hospital Cambridge) whose
 assistance we gratefully acknowledge.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/
 Location/Qualifiers
 1. 2637
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NOD"
 /db_xref="FANTOM:DB:E43J0023D24"
 /db_xref="taxon:10090"
 FEATURES
 source

CDS

```

/clone="E43002D24"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
358..1137
/notes="unnamed protein product; MODULATOR OF ANTIGEN
RECEPTOR SIGNALING MARS (SPRR AL18196, evidence: PASTY,
100%ID, 100%length, match=777)
putative"
/codon_start=1
/protein_id="BAC04095.1"
/db_xref="GI:26353730"
/translation="MGLSSRGKTSPPSSSSGPDQPEVSNQPERHKYVALGSLFPA
GDPARSLIEGPRITLISDDGMVTQVSLSREYHMSVYAKVHQLVEGLSREK
ASBLULPKNPGGAFILRSQTRGCYSLSVPLSPASMDLIRHRIQLRNLGMLYIS
PRTPEPSHALVEHSLADGICCPUREPCVQKIGPLPGMDTPPPVTVPSLWKK
LRLSLFLRPAASGEASLISLEGLRESLSSYSISLADPDUDA"

```

ORIGIN

Query Match	Best Local Similarity	81.9%	Score 537.8	DB 3	Length 2637
Matches 646	Conservative	0	Mismatches 137	Indels 6	Gaps 2
QY	1	ATGGGAAGTGTGCCCCAGCAGAAAGAAATCTCTGCAAGCCCAAGCTTGAGTTCTCTGTC	60		
Db	358	ATGGGAAGTTGTTCACAGAGGAAAC---CTCAGGCCCAAGCCCAAGCTCTCTGTGT	414		
QY	61	CAAGCCAGGAGCCTGTGACCATGAGACGAGAGAAACAGCCACAGCCCTGGCCCTG	120		
Db	415	CAAGACCAAGAACCCGTCGTCCATGCAACCAAGAAAGACACAGGTCAACACTGTGGCCCTG	474		
QY	121	GGCAGTTCCCGCAGAGTGGCCCGCGCACTGTGCGTGAAGTCCGAGGAGCATTGACC	180		
Db	475	GGCAGTTCCCAAGAGGTGAACAGGCCAGACTATCTCTGAGACTTCGGGAGCCGCTGACC	534		
QY	181	ATCGTCTGTGAGATGAGAGACTGGTGGACCGGTCTGTGAACTCTCAGCAGAGATAT	240		
Db	535	ATCATCTCTGAGATGAGATGGTGTGAGACAGTCAAGTCGAGACTCAGCAGAGATAC	594		
QY	241	AACATCCCAAGCTCCACGTGGCCAAAGTCTCCCATGAGTGGCTGTATGAGGGCCCTGAGC	300		
Db	595	CACATGCCCATGTGTATGTGTGCTAAAGTGGCCACCGGTGGCTGTATGAGGGCCCTGAGC	654		
QY	301	AGGGAAGAAAGCAGAGAGAACTGTGTGTACCTGTGGAAACCTGTGAGGGGACCTCTCATC	360		
Db	655	CGGAGAAAGCCGAGGAACTACTCTGTATCCTGTGAGAAACCCCGAGGGGACCTTCTCATC	714		
QY	351	CGGGAGACCCAGACAGAGAGAGCTTTACTCTGTCACTCGGCTCAGCCGCGCTTGCA	420		
Db	715	CGGAGAGCCAGACCAAGAGAGGCTGCTATTCCTGTCCGTCGCACTCAGCCGCGCTTGCA	774		
QY	421	TCTGTGGAGCGGATCAGACACTACAGATCCACTGCTTGACATATGCTGGCTGTATATC	480		
Db	775	TCTTGGAGCCGGATCAGACACTACAGATACAGGTCTTGACATATGCTGGCTGTATATC	834		
QY	481	TACCCGCCCTCACCCTTCCCTCACTCCAGGCCCTGGAGCACTTACTCTGAGCTGAGC	540		
Db	835	TACCTGCGCTCACTTCCCTCACTCCAGCCCTTGGTGAAGCACTTACTCTGAGCTGACA	894		
QY	541	GATGACATCTGTGCTACTCAAGAGAGCCGTGTCTTGAGAGAGGCTGCGCGCTCCCT	600		
Db	895	GATGGATCTGTCTTCCCTCAGAGAGCCGTGTCTTGAGAGAGGCTTGGGCTCACTACT	954		
QY	601	GGCAAGATATACCCCTACCTGTGACTGTGAGAGAGACCACTCACTACGTGAAGAGCTG	660		
Db	955	GGCAAGATATACCTCACTGTGACTGTGAGCAATCATCATTAATTGGAAGAAAGCTG	1014		
QY	661	GACAGCTCCCTCTGTTTCTTAAG---CTGCAAGGGGAGAGGTCCTCTCTGAGTGA	717		
Db	1015	GACCGAGCTCTCTGTTCTTGAAACCACTGTGAGGTGGAGAGCACTCTGTCTGAGTGA	1077		
QY	718	GGTCTCCGGAGTCCCTCAGCTTTCATCATCAGCTGAATGACGAGGCTGTCTCTTGGAT	777		

[illegible]

JOURNAL

Submitted (16-JUN-2001) Yoshinobu Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Sueniwa-cho, Tsukuba, Ibaraki, Kanagawa 230-0045, Japan [E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedic Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

1..2974

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM DB:5830437K1.0"

/db_xref="Eaxon:10090"

/clone="5830437K1.0"

/sex="male"

/tissue_type="thymus"

/clone_id="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

232..1011

/note="unnamed protein product; MODULATOR OF ANTIGEN RECEPTOR SIGNALING MARS (SPTRL) [AL38196, evidence: FASTA, 10081D, 100%length, match=777]"

putative"

/codon_start=1

/protein_id="BAC27168.1"

/db_xref="GI:26326849"

/translation="MGSLSSRGKTSPPSSSGPDPVSKQPRHHVYTAVALGSPFA
GEQARLSRLRGEPITLTSBDDWMTVSEVSGRYHMPSVYAKVAHGMVEGLSRK
AEELLLPNNPVGAFELIRSQTRGYSLEVRSLRPSWRIHRYRIORLDNGLYTT
PRLTFPSHALVHTYSLADGICCPLEPCVLQKGLPGKIDPDPVPTVPTSLNWKK
LDRSLFLTEAPASGEASLISEGLRESLSYSLSLAEDPLDA"

ORIGIN

Query Match 68.2%; Score 536.2; DB 3; Length 2974;
Best Local Similarity 81.7%; Pred. No. 1.8e-130;
Matches 645; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

1 ATGGGAAGTCTGCCAGCAGAGAAGAAATCTCTGCCAAGCCCAAGCTTGAGTCTCTGTCTG 60
232 ATGGGAAGTCTGTCTCAACAAGGGAAGAAAC--CTCCAGCCCCAGCCCCAGCTCTCTGT 288
61 CAAGGCCAGGAGACCTGTGACCATGAGAGCAGAGAGAAAGCACAAGCCAGCCTGGACCTG 120
289 CCAGACCAAGAACCCCGTGTCCATGACAACCAAGAAAGACACAAGGTCAAGCTGTGGCCCTG 348
121 GGCAGTTTCCCGGAGGTGAGCCCGGCCGAGCTGTGCTGAAGTCTGGGAGACCATTGACC 180
349 GGCAGTTTCCACAGAGGTGAACAGGCCAGCACTAATCTCTGAGACTCGGGAGACCGCTGACC 408
181 ATGCTCTCTAGAGTGTGAGACTGCTGTGAGACGCTGCTCTGAATCTTCAGGCCACAGAGTAT 240
409 ATATCTCTTAGAGATGAGATTGATGAGACAGTCCAGTCGGAATCTTCAGGCAGAGATAC 468
241 AACATCCCAAGCGTCCACGTTGACCAAGTCTCCCATGAGTGTGCTGATAGAGGACCTGAGC 300
469 CACATGCCCGATGTGATGTATGTGCTAAAGTGGCCACAGGATGCTGTATCAGAGGCCCTGAGC 528
301 AGGAGAAAGACAGAGAACTGCTGTTGTTAATCTTGAGAAACCTTGAGAGGGGCTTTCTCATC 360
529 CGGAGAAAGACCGAGAAACTACTCTCTGTTAATCTTGAGAAACCCCGAGAGGGGCTTCTCATC 588
361 CGGAGAGCGCAGACCGAGAGAGGCTCTTACTCTCTGTCAAGTCCGCTCAGCGCCCTCGCA 420
589 CGGAGAGCGCAGACCGAGAGAGGCTACTATCTCTCTGTCTGTCTCGCATTCAGCGCCCTCGCA 648

QY 421 TCCTGGGACCGGATCAGACACTACAGATCACTGCTTGAATGGTGGTATCATC 480
 DB 649 TCTTGGGACCGGATCAGACACTACAGATCACTGCTTGAATGGTGGTATCATC 708
 QY 481 TCACCGCGGCTCACTTCCCTCTACTCCAGGCGCTGGTGGACCTTACTGTGAGCTGGG 540
 DB 709 AACACCTCGCTCACTTCCCTCTACTCCAGGCGCTGGTGGACCTTACTGTGAGCTGGG 768
 QY 541 GATGACATCTGCTGCTTACTCAGAGAGCCCTGTGCTTGAAGGGCTGGCCGCTCCCT 600
 DB 769 GATGACATCTGCTGCTTACTCAGAGAGCCCTGTGCTTGAAGGGCTGGCCGCTCCCT 828
 QY 601 GGCAGATATATACCCCTTACTGCTGAGAGAGACACCTCACTCACTGAGAAAGCTG 660
 DB 829 GGCAGATATATACCCCTTACTGCTGAGAGAGACACCTCACTCACTGAGAAAGCTG 888
 QY 661 GACAGCTCCCTCTGTTTCTGAAAG--CTGCCAGAGGAGAGTCTCTTCACTGAG 717
 DB 889 GACCGAGCTCTGTTTCTGAAAGCACTGCGAGTGGGAGGCACTCTGCTCAGTGA 948
 QY 718 GGTCTCCGAGAGTCTCTGCTTACTGATGAGCTGAGTGAAGAGCTCTCTTGGAT 777
 DB 949 GGGCTCCGAGAGTCTCTGCTTACTGATGAGCTGAGTGAAGAGCTCTCTTGGAT 1008
 QY 778 GATGCTAG 786
 DB 1009 TAGCCTGG 1017

RESULT 6
 AL541041 1032 bp mRNA linear EST 24-MAR-2004
 LOCUS
 DEFINITION
 AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23
 5-PRIME, mRNA sequence.
 ACCESSION
 AL541041
 VERSION
 AL541041.3 GI:45716635
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS
 TITL
 JOURNAL
 COMMENT
 On Feb 15, 2001 this sequence version replaced gi:30544829.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(4T) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 9825.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna/s=CS0DE005AF12QP1&c=9825.r.
 Location/Qualifiers
 1..1032
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE005YK23"
 /issue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(4T) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

FEATURES
 source
 .ORIGIN

Query Match 67.3%; Score 529; DB 1; Length 1032;
 Best Local Similarity 96.5%; Pred. No. 1,1e-128;
 Matches 526; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGGAAGTCTGCTCCCAAGCAAGAAATCTCTGCGCAAGCCCAAGTGTGAGTCTCTGTC 60
 DB 395 ATGGGAAGTCTGCTCCCAAGCAAGAAATCTCTGCGCAAGCCCAAGTGTGAGTCTCTGTC 454
 QY 61 CAAGCCAGAGGACCTGTGACATGGAAGCAAGAGAGCAAGGCCACAGCCGTGGCCCTG 120
 DB 455 CAAGCCAGAGGACCTGTGACATGGAAGCAAGAGAGCAAGGCCACAGCCGTGGCCCTG 514
 QY 121 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCCGTGAGATTCGGGGAGCCATTGACC 180
 DB 515 GGCAGTTTCCCGCAGGTGGCCCGCCGAGCTGTCCGTGAGATTCGGGGAGCCATTGACC 574
 QY 181 ATGCTCTGAGAGATGGAAGCTGTTGGAAGGAGGCTGTGAAATCTCAGGACAGAGTAT 240
 DB 575 ATGCTCTGAGAGATGGAAGCTGTTGGAAGGAGGCTGTGAAATCTCAGGACAGAGTAT 634
 QY 241 AACATCCGACGCTGCCACGTGGCCAAAGTCTCCATGGGTGATGAGGAGCTGAGC 300
 DB 635 AACATCCGACGCTGCCACGTGGCCAAAGTCTCCATGGGTGATGAGGAGCTGAGC 694
 QY 301 AGGAGAAAGCAGAGGAACTGCTGTTTACTCTGGGAAACCTTGGAGGGGCTTCTCATC 360
 DB 695 AGGAGAAAGCAGAGGAACTGCTGTTTACTCTGGGAAACCTTGGAGGGGCTTCTCATC 754
 QY 361 CGGAGAGCCGACGAGGAGAGGAGGCTCTTACTCTGTCAGTCCGAGCTCAGCCGCTGCA 420
 DB 755 CGGAGAGCCGACGAGGAGAGGAGGCTCTTACTCTGTCAGTCCGAGCTCAGCCGCTGCA 814
 QY 421 TCCTGGGACCGGATCAGACACTACAGATCACTGCTTGAATGGTGGTATCATC 480
 DB 815 TCTTGGGACCGGATCAGACACTACAGATCACTGCTTGAATGGTGGTATCATC 874
 QY 481 TCACCGCGGCTCACTTCCCTCTACTCCAGGCGCTGGTGGACCTTCTGAGAGCTGGCG 540
 DB 875 TCACCGCGGCTCACTTCCCTCTACTCCAGGCGCTGGTGGACCTTCTGAGAGCTGGCG 934
 QY 541 GATGA 545
 DB 935 RTKAA 939

RESULT 7
 AK020837 926 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION
 AK020837 Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 clone:A930009E21 product:MODULATOR OF ANTIGEN RECEPTOR SIGNALING
 MARS, full insert sequence.
 ACCESSION
 AK020837
 VERSION
 AK020837.1 GI:12861542
 KEYWORDS
 HTC; CAP tripper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1
 AUTHORS
 TITL
 JOURNAL
 MEDLINE
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 MEDLINE
 PUBMED
 11042159

REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
TITLE	Riken integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multichipillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	20530913
REFERENCE	
AUTHORS	11076861
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 926)
TITLE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaya, S., Kuwahara, C., Kawai, J., Kojima, Y., Komono, H., Kotaka, M., Nomura, K., Numazaki, R., Ono, M., Okazaki, A., Nishi, K., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sato, T., Owa, C., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details.
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAAGAGACGATCCAGACTCTTTTTCCTTTTTCCTTA 3'], cDNA was prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAAGATTCGAGTTAATTAAATCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified phagescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Guscinich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.
TITLE	Retina RNA was provided by Stefano Guscinich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.
JOURNAL	Location/Qualifiers
FEATURES	1..926
SOURCE	/organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J"

Query Match	Best Local Similarity	83.1%	Score 478;	DB 3;	Length 926;
Matches 557;	Conservative 0;	Mismatches 110;	Indels 3;	Gaps 1;	
QY 120	GGGCAAGTTCCCGCAGGTGGCCCGCCGACGCTGCTGACTCGGGAGCCGANTTAC	179			
DB 4	GGGCAAGTTCCCGCAGGTGGAAAGGCGCAACTATCTTAAGCTCGGGAGCGCGCTGAC	63			
QY 180	CATGCTCTGAGGATGAGAGTGTGTGACGAGTGTCTGGAAGTCTCAGGACAGAGTA	239			
DB 64	CATCATCTCTGAGATGAGATTGTGTGACAGTCCAGTCCGAAAGTCTCAGGACAGAGTA	123			
QY 240	TAAATCCCCAGGCTCAAGTGGCCAAAGTCTCCATGGTGGCTGTATGAGGAGCTGAG	299			
DB 124	CCACATGCGCAAGTGTGTATGTGGTAAAGTCCGCGCGGTGGCTGTATGAGGAGCTGAG	183			
QY 300	CAGGAGAAAGCAGAGAACTGCGTGTGTTAAGTGGGAAACCTGGAGAGGCGCTTCCGAT	359			
DB 184	CCGGAGAAAGCCGAGAACTACTCTGTGTTAAGTGGGAAACCTGGAGAGGCGCTTCCGAT	243			
QY 360	CCGGAGAGCCCAACACAGAGAGGCTTTAATCTCTGTGAGTCCGCTCAGCCGCGCTGC	419			
DB 244	CCGGAGAGCCCAACACAGAGAGGCTTTAATCTCTGTGAGTCCGCTCAGCCGCGCTGC	303			
QY 420	ATCTGAGGACCGATTCAGACATTCAGAGATCCACTGCTTGAACATGGCTGTATCAT	479			
DB 304	ATCTGAGGACCGATTCAGACATTCAGAGATCCACTGCTTGAACATGGCTGTATCAT	363			
QY 480	CTCACCGCGCTCACTCTCCCTCAGTCCAGGCGCTGTGAGCACTTACTCTGAGCTGAC	539			
DB 364	CTCACCTGCGCTCACTCTCCCTCAGTCCAGGCGCTGTGAGCACTTACTCTGAGCTGAC	423			
QY 540	GGATGACATCTGCTGCTTACTCAAGAGAGCCCTGTGTCTTGACAGAGGCTGCGCCGCTCC	599			
DB 424	AGATGGGATGTGCTGTCCCTCAGGAGAGCGGTGTGTCTTGACAGAACTTGGGCGCACTTAC	483			
QY 600	TGGCAAGATATACCCCTACTGTGACTGTGACAGAGACCACTCACTGAGAAAGAGCT	659			
DB 484	TGGCAAGATATACCCCTACTGTGACTGTGACAGAGACCACTCACTGAGAAAGAGCT	543			
QY 660	GGACAGCTCCCTCTGTTTCTGAAG--CTGCAACAGGAGAGAGTCTTCTCAGTGA	716			
DB 544	GGACCGGAGGCTCTGTTTCTGAAACACCTGCGAGATGGGAGAGGCACTCTGCTCAGTGA	603			
QY 717	GGGCTCCGAGAGCTCAGCTTCAATCAAGCTGAAATGACAGAGCTGTCTTTGGA	776			
DB 604	GGGCTCCGAGAGCTCAGCTTCTTAATCAAGCTGTGAGAGACCTCTTGAGATGATGC	663			
QY 777	TGATGCTTAC	786			

QY 484 CTCACCGGCGCTCACCTTCCCTCACTCCAGAGCGCCGTGGACCATTAATCTCTAGCTGGC 539

Db 364 CTCACCTGCGCTCACCTTCCCTCACTCCAGCTTGGAGCAATTACTCTGAGCTAGC 423

QY 540 GSATGACATCTGTGCTACTCAAGAGCCCTGTGTCTGCAAGAGGAGCTGGCCGCTCC 529

Db 424 AGATGGCATCTGTGCTGCCCTCAGGAGCGGTGTGTCTGCAAGACCTGGGCATACC 483

QY 600 TGGCAAGGATATACCCCTCACTGTGACTGTGAGAGAGAACCACTCACTGGAAAGAGT 659

Db 484 TGGCAAAAGTATACCTTCCACCTGTGACTGTGCAACATATCACTAAATTGGAAAAAGCT 543

QY 660 GSAACAGTCCCTCTGTTTTCTGAAG---CTGCCACAGGGGAGAGTCTCTTCTCAGTGA 716

Db 544 GGACCGTAGCCCTCTGTTTTCTGGAAGACCTGCAAGTGGGGAGGCATCTGTGATTTGA 603

QY 717 GGGTCTCCGGGAGTCCCTCAGCTTCTACATCAGAGCTGAATGACGA 761

Db 604 GGGGCTTCAGAGT-CTTCACCTTTCTACTCTCAATCCGGATTGATGA 647

RESULT 9			
LOCUS	BC284179		
DEFINITION	BC284179	566 bp	mRNA
ACCESSION	U024082261	NIH_MGC_91	Homo sapiens cDNA clone IMAGE:4520382 5',
VERSION	BC284179		
KEYWORDS	BC284179.1	GI:13034866	
SOURCE	EST.		
ORGANISM	Homo sapiens		
	Homo sapiens (human)		

REFERENCE	(bases 1 to 566)
AUTHORS	NIH-MGC http://mgc.ncbi.nlm.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

```
found through the I.M.A.G.E. Consortium/ILINK at:
http://image.jnl.gov
plate: ILINK0418 row: C column: 07
High quality sequence start: 2
High quality sequence stop: 566.
Location/Qualifiers
1..566
FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4520382"
/tissue_type="adenocarcinoma, cell_line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

Query Match	51.5%	Score 405	DB 4	Length 566
Best Local Similarity	100.0%	Pred. No.	100-96	
Matches 405	Conservative 0	Mismatches 0	Indels 0	Gaps 0
382	GGCTCTACCTCTGTGTAGACCGCGCTAGCGCCCTTGATCTTGGACCGGATAGACAC	441		
13	GGCTCTTACCTCTGTGTAGTCGCGCTAGCGCCCTTGATCTTGGACCGGATAGACAC	72		

QY	442	AACAGGATCCATCGCTTTGAACAATGGCGGCTGTACATTCACCGCGGCTACCTTTCCC	501
Db	73	TACAGGATCCATCGCTTTGAACAATGGCGGCTGTACATTCACCGCGGCTACCTTTCCC	132
QY	502	TCATCCAGGCGCTGTGGACCATTTACTCTGTAGCTGGGCGGATGATCATCTGCGCTACTTC	561
Db	133	TCATCCAGGCGCTGTGGACCATTTACTCTGTAGCTGGGCGGATGATCATCTGCGCTACTTC	192
QY	562	AAGGAGCCCTGTGTCTCTGACAGAGGCTTGCCCGCTCCCTGGCAAGATATATACCCCTACTC	621
Db	193	AAGGAGCCCTGTGTCTCTGACAGAGGCTTGCCCGCTCCCTGGCAAGATATATACCCCTACTC	252
QY	622	GTGACTGTGCGAGAGCACCACTCAATGTGAAAAGACTGGACACTCCCTCTCTGTTTCT	681
Db	253	GTGACTGTGCGAGAGCACCACTCAATGTGAAAAGACTGGACACTCCCTCTCTGTTTCT	312
QY	682	GAGCTGCGCAAGGGGAGAGGCTCTTCTTCAGTAGGGTCTTCGGGAGTCCCTCAAGCTTC	741
Db	313	GAGCTGCGCAAGGGGAGAGGCTCTTCTTCTTCAGTAGGGTCTTCGGGAGTCCCTCAAGCTTC	372
QY	742	TACATCAGCCCTGATATGACGAGGCTGTCTCTTTGGATGATGCTGAG	786
Db	373	TACATCAGCCCTGATATGACGAGGCTGTCTCTTTGGATGATGCTGAG	417

RESULT	10
B0054265	
LOCUS	
DEFINITION	B0054265 986 bp mRNA linear EST 29-MAR-2002 AGNCCOURT_6630248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
ACCESSION	B0054265
VERSION	B0054265
KEYWORDS	B0054265.1 GI:19813605
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 986)				
	NIH-MGC	http://mgc.ncl.nih.gov/		
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.			

```

FEATURES
source
1. 986
location/Qualifiers
1. 986
plate: ILCM2125 row: 1 column: 12
high quality sequence stop: 515.

```

```

/organism="Homo sapiens"
/mol_type="tRNA"
/db_xref="taxon:9606"
/clone_image:5936339"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_106"
/note="Organ: blood; Vector: pOT7; Site1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally clone into EcoRI/XhoI sites using the following 5' adaptor: GGACGAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN
Query Match      Score 386.2,  DB 5,  Length 986;
Best Local Similarity  98.0%;  Pred. No. 6,3e-91;

```


ORIGIN

/clone_1lb="NIH_MGC_107"
 /note="Organ: breast; Vector: pOTB7; Site: 1: EcoRI;
 Site: 2: XhoI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."

Query Match 40.8%; Score 321; DB 5; Length 960;
 Best Local Similarity 88.4%; Pred. No. 1,1e-73;
 Matches 381; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

QY 356 TCATCCGGGAGAGCCAGACAGAGAGGCTTACTCTGTGATGCGCTGAGCGCC 415
 DB 1 TCATCCGGGAGAGCCAGACAGAGAGGCTTACTCTGTGATGCGCTGAGCGCC 60
 QY 416 CTGCATCTGGGAGCCGAGATGAGACACTAGAGATCCATGCTTGAATGGCTGCT 475
 DB 61 CTGCATCTGGGAGCCGAGATGAGACACTAGAGATCCATGCTTGAATGGCTGCT 120
 QY 476 ACATCTACCGGCGCTGACCTTCCCTCACTCCAGGCGCTGTGAGCATTTACTGAGC 535
 DB 121 ACATCTACCGGCGCTGACCTTCCCTCACTCCAGGCGCTGTGAGCATTTACTCT 176
 QY 536 TGGCGGATGACATCTGCTGCTTACTCAAGAGCCCTGTCTGAGAGGAGGCTGCGCCG 595
 DB 177 -----GAGGCGCTGCGCCG 190
 QY 596 TCCTCGGAGAGATATACCTTACTGATGCTGAGAGAGACCACTCACTGAGAG 655
 DB 191 TCCTCGGAGAGATATACCTTACTGATGCTGAGAGAGACCACTCACTGAGAG 250
 QY 656 AGCTGAGAGCTCCCTCTGTTTCTGAAGCTGACAGAGGAGAGCTCTTCTCAGTG 715
 DB 251 AGCTGAGAGCTCCCTCTGTTTCTGAAGCTGACAGAGGAGAGCTCTTCTCAGTG 310
 QY 716 AGGCTTCGGGAGTCCCTCAGCTTCTTACTGATGAGCTGATGAGAGGCTCTCTTGG 775
 DB 311 AGGCTTCGGGAGTCCCTCAGCTTCTTACTGATGAGCTGATGAGAGGCTCTCTTGG 370
 QY 776 ATGATGCTTAG 786
 DB 371 ATGATGCTTAG 381

RESULT 15

AL844307 614 bp mRNA linear EST 30-JUL-2002
 DEFINITION AL844307 pool_AK_1lb_v_SPD Homo sapiens cDNA, mRNA sequence.
 ACCESSION AL844307
 VERSION AL844307.1 GI:22019089

KEYWORDS

EST.
 Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 614)
 Ashcroft, K., Bethel, G., Bye, J.M., Howell, G.R., Huckle, E.J. and
 Sheridan, E.

TITLE Homo sapiens EST sequence
 JOURNAL Unpublished (2002)
 COMMENT The Sanger Centre
 The Sanger Centre
 Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: humguery@sanger.ac.uk
 Sanger Centre name: scod10816.154136A
 Homo sapiens EST sequence. This sequence was generated as part of
 The Wellcome Trust Sanger Institute program to identify and
 annotate genes in the human genome. Incomplete or unconfirmed genes
 are experimentally analysed using a variety of cDNA library

resources. This sequence was obtained from a PCR product generated
 from a pool of up to 100,000 cDNA clones derived from
 pool AK_1lb v SPD cDNA library. Further information can be found at
<http://www.sanger.ac.uk/Teams/Team9/>.
 Location/Qualifiers

1..614
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="20"
 /clone_1lb="pool_AK_1lb_v_SPD"

ORIGIN

Query Match 40.4%; Score 317.4; DB 1; Length 614;
 Best Local Similarity 99.7%; Pred. No. 8.8e-73;
 Matches 318; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 468 CTGGCTGTACATTCACCGGCTCAACCTTCCCTCACTCCAGGCGCTGTGAGACCATTA 527
 DB 1 CTGGCTGTACATTCACCGGCTCAACCTTCCCTCACTCCAGGCGCTGTGAGACCATTA 60
 QY 528 CTCTGAGCTGGGAGAGATGACATCTGCTGCTTCAAGAGCCCTGTGTCTGAGAGGCG 587
 DB 61 CTCTGAGCTGGGAGAGATGACATCTGCTGCTTCAAGAGCCCTGTGTCTGAGAGGCG 120
 QY 588 TGGCCGCTCCCTGAGAGAGATATACCTTACTGATGCTGAGAGAGACACTCA 647
 DB 121 TGGCCGCTCCCTGAGAGAGATATACCTTACTGATGCTGAGAGAGACACTCA 180
 QY 648 CTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGAGAGAGTCTCT 707
 DB 181 CTGGAAGAGCTGACAGCTCCCTCTGTTTCTGAAGCTGCCACAGGAGAGAGTCTCT 240
 QY 708 TCTCAGTGAAGGCTCTCCGAGAGTCCCTCAGCTTCTACATGAGCTGATGAGAGGCTGT 767
 DB 241 TCTCAGTGAAGGCTCTCCGAGAGTCCCTCAGCTTCTACATGAGCTGATGAGAGGCTGT 300
 QY 768 CTCTTTGATGATGCTTAG 786
 DB 301 CTCTTTGATGATGCTTAG 319

Search completed: November 16, 2004, 23:30:37
 Job time : 3256 secs